Studies in Second Language Learning and Teaching
Department of English Studies, Faculty of Pedagogy and Fine Arts, Adam Mickiewicz University, Kalisz
SSLLT 12 (4). 2022. 697-719
https://doi.org/10.14746/ssllt.2022.12.4.8
http://pressto.amu.edu.pl/index.php/ssllt

Models as written corrective feedback: Effects on young L2 learners' fluency in digital writing from product and process perspectives

Raquel Criado
University of Murcia, Spain
https://orcid.org/0000-0003-0444-3901
rcriado@um.es

Aitor Garcés-Manzanera
University of Murcia, Spain
https://orcid.org/0000-0002-1789-9046
aitor.garces@um.es

Luke Plonsky
Northern Arizona University, Flagstaff, USA
https://orcid.org/0000-0002-5791-1839
luke.plonsky@nau.edu

Abstract

This study was motivated by Truscott's (1996, 2004) scarcely empirically tested claims that written corrective feedback (WCF) processing hinders fluency in subsequent rewriting owing to learners' purposeful avoidance of making mistakes by composing shorter texts at a higher speed. It examined the writing fluency of the texts produced by eighteen 10-11-year-old L2 English children in a digital environment. They were divided into a feedback (*N* = 10) and a self-correction group (*N* = 8). Both groups engaged in a three-stage task: writing, comparison of their texts with a model or self-editing as appropriate, and rewriting. Fluency was analyzed via five product/offline and five process/online measures. The texts and writing behaviors were recorded with Inputlog 8.0. The results partially support Truscott's claims. The feedback group improved their fluency in all the ten measures. However, the self-editing group showed higher fluency than the feedback group in seven of the ten measures, with the corresponding Hedge's effect sizes between groups ranging from small to large. The study enlightens our knowledge of young learners' writing fluency and supports adopting a multidimensional approach to understand the complex and multi-faceted nature of fluency as mediated by WCF processing.

Keywords: written corrective feedback; models; fluency; L2 writing; young learners

1. Introduction

In 1996, Truscott claimed that grammar correction was ineffective and harmful since "students shorten and simplify their writing in order to avoid corrections" (p. 355). To support his controversial claim, he cited Semke's (1984) study, in which fluency was measured as the number of words written in a second language (L2) ten-minute timed task. The texts written by the group who were provided with content comments were significantly more fluent than those produced by the groups who received indirect feedback, direct feedback, or direct feedback and content comments. In 2004, Truscott criticized Chandler's (2003) work, whose second study revealed that students increased their speed fluency and improved their accuracy in five identical essays after processing direct and indirect written corrective feedback (WCF). Truscott claimed again that such results were due to students' purposeful avoidance of making mistakes. He also asserted that such avoidance resulted in simplified texts written at a higher speed than complex ones.

Despite the abundant research on WCF (see Karim & Nassaji, 2020, for a recent review), Truscott's claims about the negative effect of WCF on fluency have virtually gone overlooked (as noted by Ekanayaka & Ellis, 2020). Hence, further investigation is justified due to the theoretical and applied relevance of broadening our understanding of the role of WCF in L2 writing. Regarding the study of fluency in writing, over the last two decades there has been an interest in a process-oriented approach to L2 writing research focused on the identification and measurement of online, directly observable behaviors of L2 writers to infer the underlying cognitive processes in writing – planning, translation or linguistic encoding, or revision (e.g., Révész et al., 2017). In this context, computer keystroke logging programs such as Inputlog unobtrusively record learners' number, length and location of pauses, deleted and inserted characters, and mouse clicks. Yet, the majority of previous digital L2 writing research has been conducted with adults (but see Garcés-Manzanera, 2021, for a study with children).

The present study attempts to contribute empirically and pedagogically to previous work by shedding light on the relationship between WCF and young L2 learners' fluency in digital writing. Fluency was considered within a multidimensional approach (Van Waes & Leijten, 2015), that is, from both perspectives of the product or final text and of the underlying composing processes. As the WCF technique, we selected models, which have often been used with child participants in recent years (Coyle et al., 2018; Lázaro-Ibarrola, 2021). Empirically, the evidence obtained about young learners' fluency in digital writing would aid to develop our understanding of feedback processing by focusing on a population hitherto unexplored. Pedagogically, such an understanding would help teachers' decision making concerning the implementation of models as WCF when teaching young learners.

2. Literature review

2.1. Conceptualizing writing fluency: Product and process perspectives

There is theoretical and empirical consensus that complexity, accuracy, and fluency (CAF) represent the fundamental dimensions of the constructs of L2 proficiency and performance (Housen & Kuiken, 2009). In their classical review of written CAF measures, Wolfe-Quintero et al. (1998) contributed three widespread perspectives to the definition of writing fluency, namely: temporal, speed-based, and cognitive. Regarding the temporal dimension, they stated that fluency is "a measure of the sheer number of words or structural units a writer is able to include in their writing within a particular period of time" (p. 14). The same authors added speed as a key characteristic of fluency, which they defined as "the rapid production of language" (p. 117). For Wolfe Quintero et al., fluency is mostly related to the learners' control in accessing their current language knowledge. Thus, as a cognitive construct, fluency entails the "efficient access to a rich linguistic knowledge base and the (equally efficient) retrieval of propositions for utterances" (Van Gelderen & Oostdam, 2002, p. 241) that are necessary in the encoding of ideas that will be converted to linguistic forms.

The temporal and speed-based perspectives are helpful to describe fluency in the final written product. However, they do not refer to the writing actions and processes that contribute to the generation of that final text. In this respect, we will draw on Kellogg's model of cognitive processes of L1 writing (Kellogg, 1996), similar to most previous studies (e.g., Michel et al., 2020; Mohsen, 2021; Révész et al., 2017). Kellogg expanded Hayes and Flower's (1980) classical three-process model of: 1) planning ideas, 2) translation of such ideas into written sentences (through grammatical, phonological, and orthographic processing) and 3) the monitoring of ideas and of previously produced text. Kellogg subdivided monitoring into reading and editing and added two motor processes: programming and execution. Similar to Hayes and Flower, he conceived the overall process of writing as the sum of recursive and interactive operations of planning, sentence generation, and idea and text monitoring. Kellogg's model also specified the demands of the writing processes on the three components of working memory distinguished by Baddeley (1986): the central executive, phonological loop, and visuo-spatial sketchpad.

In Kellogg's (1996) model, fluency in writing is affected in different ways (regardless of the time available for the writing task), all of which are mediated by the limited resources of working memory. For example, concerning inexperienced L2 writers, whose profile corresponds to that of the participants in this study, lack of language proficiency is a crucial factor (Abdel Latif, 2009; Palviainen et al., 2012). Since they have not sufficiently proceduralized their language knowledge, their attentional resources will tend to focus on local narrow problems (spelling, grammar, etc.) at the expense of higher-level text features (discourse, cohesion, coherence, etc.), speed of production (fluency), and other writing processes such as revision and planning (Schoonen et al., 2009).

2.2. Analyzing writing fluency: Product and process measures

The consideration of fluency as one dimension of the CAF tripartite framework has resulted in writing fluency measures focused on the final outcome or product. For instance, Kellogg's (1990) pioneering study analyzed fluency by using the total number of words in timed writing, total number of words per minute in composing/execution time (without initial planning time), and total number of words in total writing time (composing/execution time plus initial planning time). In contrast, Wolfe-Quintero et al. (1998) concluded that the most reliable written CAF measures were based on the length of production units: T-unit length, error-free T-unit length, and clause length.

However, relying on all these static, offline, text-based product measures to portray fluency has been criticized given that they do not consider the "real-time, on-line production processes" (Stevenson, 2005, p. 135) that allow writers to devise a text more fluently (Chenoweth & Hayes, 2001). Thus, a multidimensional approach to studying L2 writing fluency, considering both product and process measures, is preferable (Van Waes & Leijten, 2015). For example, Chenoweth and Hayes (2001) operationalized fluency in the traditional way, that is, as words per minute. They also included pause and revision bursts, which respectively measure the number of words occurring between pauses and revisions.

Previous research (Chenoweth & Hayes, 2001; Mohsen, 2021; Palviainen et al., 2012; Van Waes & Leijten, 2015, etc.) has characterized fluent writing from process and product perspectives. L1 and L2 fluent writers engage in fewer pauses at linguistic boundaries (words and clauses) and their pausing time is more reduced. Fluent writers also make use of fewer large revisions focused on low-order processes (such as lexical selection, grammar or spelling choices), but they deal with more extensive revisions for discourse-based elements. Finally, fluent writers also display a higher production rate.

In L2 digital writing, fluency has been examined primarily with adults and in relation to text quality and task-related factors. Task complexity operationalized as the provision versus no provision of content (ideas) did not involve any significant effects on fluency (Révész et al., 2017), with mixed findings in the case of directing learners to include content from source materials (integrated tasks) or making them use their own resources (independent tasks; e.g., Barkaoui, 2019; Michel et al., 2020). Different indices of fluency have been found to be predictors of text quality (burst and fluency during burst in Spelman Miller et al., 2008; characters per P-burst and P-burst processing time in Mohsen, 2021, etc.). Also, the effects of two planning techniques (i.e., note-taking versus freewriting) on L1 German and L2 English writing fluency were examined in Breuer's (2019) small-scale study. The descriptive tendencies showed that freewriting enhanced the efficacy of writing fluency for both languages.

2.3. Models as WCF: Theoretical and empirical support

Models are native or native-like texts adapted to the learners' age and proficiency level (Coyle et al., 2018). The cognitive rationale for using models as a feedback strategy is linked to Schmidt's noticing hypothesis (2001) and Swain's output hypothesis (1995). When learners produce oral or written output in the target language, they can notice gaps in their interlanguages (a requisite for L2 learning, according to the noticing hypothesis). This noticing acts a "primer" which prepares them to optimize their subsequent use of models. By comparing the model with their own compositions, L2 writers might engage in a deeper cognitive comparison of the vocabulary, grammar, spelling, discursive and content features of both texts than that resulting from reading highlighted and/or errors corrected by their teachers (Coyle & Roca de Larios, 2020).

Past research with models has examined adults (Hanaoka, 2007), adolescents (García-Mayo & Labandibar, 2017; Kang, 2020; Martínez-Esteban & Roca de Larios, 2010), and children (Cánovas Guirao, 2018; Cánovas Guirao et al., 2015; Coyle et al., 2018; Lázaro-Ibarrola, 2021; Luquin & García-Mayo, 2021; Roothooft et al., 2022). Models have been implemented in a three-stage writing sequence: 1) a writing stage where participants are also pushed to notice the linguistic features they cannot express or have difficulties with (problematic features noticed); 2) a comparison stage of their own texts with a model text; and 3) a rewriting stage where participants try to include the features noticed in the previous stage into their own texts. A general finding from the previous body of work is that participants tend to focus on lexical and content items. They also manage to incorporate a reasonable number of features noticed into their final drafts, which become more accurate and cohesive than their initial ones and those of the learners who self-edited their own texts without the aid of models.

Several reasons account for selecting models to examine the effects of WCF on fluency in this study. First, as previously seen, recent WCF works using this technique have often focused on children, the same age group as that of our participants. Second, as shown in the aforementioned results of past research, models have emerged as a promising, student-centered and discourse-based alternative to traditional, teacher-guided options focused on isolated errors such as direct and indirect WCF (Lázaro-Ibarrola, 2021). Third, from a pedagogical angle, offering a single model text to a whole class is considerably more manageable to handle by a teacher than the time-consuming process of offering individual direct or indirect WCF to each student (Coyle & Roca de Larios, 2020). Fourth, from a cognitive perspective, using models to test their effects on fluency could be justified by parallelism with Skehan's (2009) application of Levelt's (1989) model of L1 speech production to L2 CAF oral performance. As stated above, contrary to direct and indirect WCF, models offer a large array of lexical-, mechanics-, syntactic- and discursive-related alternatives to their own original choices in the initial texts that students are pushed to notice in the comparison or WCF processing phase, and which, ideally, they should incorporate as much as possible in their rewritten texts. The effects of processing models could arguably resemble those established by Skehan for a planning stage. Basically, it is expected that the priming of the lexical features in the comparison phase would ease the pressure exercised on Kellogg's (1996) translation process, or Levelt's (1989) formulator stage for speaking in the rewriting session. Like a planning stage in oral language, such priming would likely allow the lexical forms to be more easily and effectively retrieved from long-term memory to working memory in the rewriting phase. Hence, Kellogg's execution process (or Levelt's articulator stage) would seem to benefit too. Thus, processing models would tend to enhance or, at least, not harm learners' fluency, which renders them a potentially suitable WCF technique to be applied in the current study.

2.4. The relationship between WCF and writing fluency: Empirical evidence

Several studies have investigated the effects of WCF on learners' fluency, viewed from a product perspective in hand-written texts. Only two studies explicitly alluded to Truscott's (1996, 2004) claims (Chandler, 2003, and Ekanayaka & Ellis, 2020). Young learners were examined in Cánovas Guirao (2018), Lázaro-Ibarrola

(2021) and Roothooft et al. (2022). Chandler (2003), Ekanayaka and Ellis (2020) and Kim et al. (2022) studied university students, while Martínez-Esteban and Roca de Larios (2010) investigated secondary-school students, and Sánchez (2019) examined both university and secondary-school students. Individual and collaborative writing were studied in Lázaro-Ibarrola (2021) as well as Martínez-Esteban and Roca de Larios (2010), while collaborative writing alone was analyzed in Cánovas Guirao (2018) and Kim et al. (2022). The remaining studies focused on individual writing. Chandler's (2003) first reported study resorted to indirect WCF, Ekanayaka and Ellis (2020) used direct, semi-focused WCF, synchronous indirect WCF was selected by Kim et al. (2022), while Chandler's (2003) second study and Sánchez (2019) drew upon both direct and indirect WCF. Cánovas Guirao (2018), Lázaro-Ibarrola (2021) and Martínez-Esteban and Roca de Larios (2010) used models, which were also employed by Roothooft et al. (2022) together with direct WCF. Crucial methodological issues in these studies are worth mentioning:

1. Only four works included control or comparison groups (study 1 in Chandler, 2003; Ekanayaka & Ellis, 2020; Kim et al., 2022; Sánchez, 2019).
2. The sample sizes were generally small, ranging from five individual writers and six pairs in Martínez-Esteban and Roca de Larios (2010) to Ekanayaka and Ellis' (2020) thirty participants in one WCF group and thirty-one writers in both the control and the second WCF groups. Six studies reported inferential statistics (Chandler, 2003; Ekanayaka & Ellis, 2020; Kim et al., 2022; Lázaro-Ibarrola, 2021; Roothooft et al., 2022; Sánchez, 2019), whose results may not be stable due to the small samples that were used.
3. Importantly, the measurement of the fluency of the final text greatly differed: a) frequency counts, either the raw number of words in untimed collaborative writing (Cánovas Guirao, 2018), or in timed writing such as total number of words in ten minutes (Ekanayaka & Ellis, 2020), or in twenty minutes (Roothooft et al., 2022); also, the number of T-units in 25 minutes (Kim et al., 2022) and total number of words, the proportion of T-units and number of clauses per text in 30 and 20 minutes respectively for the collaborative and the individual groups (Lázaro-Ibarrola, 2021); b) temporal ratios such as words and syllables per minute in untimed writing (Sánchez, 2019); c) a speed perspective with minutes per 100 words (Chandler, 2003); and d) description of mean writing minutes and of words written (Martínez-Esteban & Roca de Larios, 2010).

A complex and inconclusive picture of the effects of WCF on fluency from a product perspective emerges from previous studies, which generally seem to support Truscott's (1996) claim for word count regardless of the type of feedback.

Kim et al.'s (2022) results pointed to a negative effect of synchronous indirect WCF on the number of T-units. Conversely, in terms of descriptive statistics, Ekanayaka and Ellis (2020) reported that the two WCF groups (with and without revision) and the control group augmented their fluency from the first to the second task. A detrimental effect of models was found for text length in Lázaro-Ibarrola's (2021) individually rewritten texts as opposed to the collaborative ones, although without statistically significant differences within groups. The same pattern emerged in the proportion of T-units and number of clauses per text. However, Roothooft et al.'s (2022) individually written texts in the model group slightly increased their word count as opposed to the direct WCF texts, without any significant differences between groups. Contrary to Lázaro-Ibarrola's (2021) results for collaborative writing, the non-model instruction group in Cánovas Guirao (2018) slightly decreased the length of their rewritten texts.

A mixed pattern appears in those studies which examined fluency with the remaining measures previously mentioned in point 3 above. After processing models, Martínez-Esteban and Roca de Larios' (2010) participants displayed higher fluency numbers in their rewritten texts, especially in collaborative writing. Sánchez's (2019) both proficiency level students who did not either receive any direct or indirect WCF or self-corrected their texts and the low-proficiency students in the self-correction group significantly improved their fluency. However, Chandler (2003) reported that all the groups in her two studies significantly increased their fluency from their first to their fifth assignments.

3. The study

3.1. Research questions

The diversity of the previous findings together with the aforementioned methodological differences in the limited number of available studies makes it difficult to reach firm conclusions about potential effects of WCF on fluency. Likewise, there are no previous studies which have jointly examined whether Truscott's claims of a negative effect of direct and indirect WCF on fluency regarded as text length (Truscott, 1996) and of increased speed fluency in writing after processing the same WCF types (Truscott, 2004) are complied with or not when the WCF technique applied is models. Finally, to the best of our knowledge, no past studies within a computer-mediated environment have examined L2 children's composing processes in general and fluency in particular. Therefore, further research on the role of WCF in young learners' writing fluency is justified and it is also warranted to look at these effects from product and process perspectives. In an attempt to fill the gaps in previous research on the effects of WCF on fluency, our study was guided by two research questions:

1) To what extent does WCF in the form of a model text affect L2 English children's writing fluency as measured by product indices?
2) To what extent does WCF in the form of a model text affect L2 English children's writing fluency as measured by process indices?

3.2. Method

3.2.1. Participants and context

The participants were 18 children who belonged to an intact 5th grade class in Spanish primary education. They were aged 10 to 11 (*M* = 10.10; *SD* = 0.31). They were attending a Spanish semi-private school where they received 3 weekly hours in the EFL subject. The teacher rated their proficiency level as not higher than the A1 level following the Council of Europe (2001). Ten children were randomly assigned to the feedback group, and 8 of them to the self-editing group. As verified by the interkey-stroke interval measure provided by Inputlog 8.0 (Leijten & Van Waes, 2013), defined as the mean number of milliseconds that happen between two characters of a word, both groups had a similar mean typing speed when composing their initial texts, with a similar high individual variability too: *M* = 210.67, *SD* = 40.69 in the feedback group and *M* = 192.27, *SD* = 39.15 in the self-editing group. The effect size between groups was small (*g* = 0.44).

3.2.2. Data collection

A one-cycle data collection procedure with three stages was implemented. In Stage 1, the children received García Hernández et al.'s (2017, p. 22) six-picture sequential prompt, which these authors had used with young learners too. Our participants were asked to type the underlying story (a scientist who tries a new potion and turns into a cat, which is attacked by the scientist's dog). Inputlog 8.0 had been installed on their computers for recording and analyzing their texts, keypresses, pausing time plus characters inserted and deleted in revisions. Due to class-time restrictions, they had a maximum of thirty minutes to complete the task. In Stage 2, twenty-four hours after Stage 1, the feedback group had access to their initial texts on the computer screen and were provided with Coyle et al.'s (2018) model text based on the six-picture story task. They were also given a prompt which directed them to write the differences between the model and their own text. The self-editing group accessed their initial texts on the computer screen too. They were given a prompt for self-editing their own compositions without the model and asked to revise their text in order to improve it. The time allotted to Stage 2 was sixty minutes, which was deemed to be a reasonable time frame to follow Cánovas Guirao's (2018) recommendation of allowing sufficient time to optimize feedback processing. Both the task and the two prompts were printed. Stage 3 took place four days after Stage 2. The two groups were asked to write the same six-picture story under identical conditions to those of Stage 1.

3.2.3. Fluency measures

Our approach to fluency was multidimensional (Van Waes & Leijten, 2015), thus covering both product and process perspectives. The selection of the fluency measures that were used in the present study addressed the following considerations:

- the wide array of measures used in previous research, with the resulting difficulty in the comparability of findings;
- the absence of any standardized criterion to discriminate among each product and process measure and/or to ensure avoidance of multicollinearity;
- the absence of any previous studies which had examined fluency from a process-oriented perspective in L2 children.

We finally opted for ten fluency measures: five product- and five process-oriented ones. We included word-based measures as they are abundant in the literature. However, we also resorted to character-based measures since they take the lengths of words into account (Palviainen et al., 2012). Out of the ten measures, six were automatically retrieved from Inputlog: word count, words per minute, characters per minute, linear fluency I, linear fluency II and product/process ratio.

Our *product* measures of fluency were as follows:

- *Word count.* According to Wolfe-Quintero et al. (1998), although computing the number of words in timed writing (as is our case; see section 3.2.2.) allows the comparison of participants' fluency within the same study, these authors maintain that "the results are meaningless in comparison with other populations or across different tasks" (p. 10). Abdel Latif (2009, 2013) adds that text quantity may depend on writers' pre-task decisions about the number of words, lines or paragraphs their text will comprise or their familiarity with the topic. We understand all these cautionary points but decided to include raw word count to test Truscott's (1996) claim that WCF harms fluency from a text-length perspective (p. 355), by resorting to models as the WCF technique.
- *Words per minute*. Abdel Latif (2009) argues that this traditional measure is not valid given that writing fewer words per minute can be due to writers' negative attitude towards writing or the reflective nature of a writer rather than lack of fluency. However, the words per minute index allows the homogenization of the measurement of fluency regardless of whether the writing task is timed or untimed and whether the students use the full time

they have available. It is calculated by dividing the total number of words produced in the final text (i.e., without including deleted words in revisions) by the total number of minutes spent on composing the written task.

- *Characters per minute*. It computes the number of characters written per minute in the final text, including spaces (Van Waes & Leijten, 2015).
- *Minutes per 20 words* (also in the final text). We included this measure to test Truscott's (2004) claim derived from Chandler's (2003) results that students write (simpler) texts at a higher speed following the processing of WCF. Accordingly, we decided to use a similar speed fluency measure to Chandler's but adapted to young learners' production. Thus, we set the threshold to 20 words (without including deleted words in revisions) after revising the word counts of our own participants, whose lowest figure was 23.
- *Minutes per 100 characters.* We set the threshold to 100 characters, since the lowest character count recorded in our data was 103. Similar to minutes per 20 words, this measure focused on the final text, including spaces, and it did not include deleted characters in revisions.

We drew upon the following *process* measures of fluency:

- *Linear fluency I: characters per minute* (labeled as "fluency (linear)" by Palviainen et al., 2012, p. 54). It takes the perspective of the linear (not final) text, which includes deleted and inserted characters plus spaces.
- *Linear fluency II: words per minute* in linear text, included by parallelism with the previous measure and which covers deleted and inserted words.

  Similar to all the product indices, following Kellogg (1996), the denominator of the two measures is the total amount of minutes spent on the composing time of the writing task (active writing time + pausing time, both in minutes).
- *Words per P-Burst*. Using burst as a unit of measure for fluency has been widespread, and it has been mostly examined from the perspective of words produced between pauses (see Chenoweth & Hayes, 2001; Mohsen, 2021; Révész et al., 2017; Stevenson, 2005). Its calculation involves the division of the total amount of typed words (including deleted ones in revisions) by the total number of P-bursts.
- *Characters per P-Burst*. It is calculated by dividing the total amount of typed characters (both deleted and inserted characters plus spaces) by the total number of P-bursts (Michel et al., 2020; Révész et al., 2017).

  According to Révész et al. (2017), both words and characters per P-burst are indices of speed fluency within a process-oriented perspective. In line with the same authors and other previous research (e.g., Mohsen, 2021), together with the absence of previous studies with children in L2 digital writing, the pause threshold was set at 2 seconds for the two measures. The number of P-bursts in both indices was automatically obtained from Inputlog.

- *Product/process ratio*. The product/process ratio is the division between the total number of characters with spaces in the final text plus the addition of the total number of non-character keys divided by the total number of characters produced in the linear text. It provides indirect information concerning how cognitively demanding the amount of revision is. The closer the ratio is to +1, the fewer revisions are undertaken, which in principle involves higher writing speed and, thus, higher fluency. We considered this measure to be relevant given our participants' age and level, since inexperienced L2 adult writers have been shown to spend considerable time on revisions, particularly of low-level aspects such as spelling and grammar errors (Barkaoui, 2007).

3.2.4. Statistical analyses

When analyzing our data, we chose to focus on descriptive statistics, including effect sizes, rather than running tests of statistical significance. This choice was based on several considerations. First, the small sample size (*N* = 18) and our interest in avoiding the increased potential for Type II errors associated with small samples. At a more conceptual level, we also chose to focus on descriptives in order to maintain an emphasis on the magnitude of the differences between the two written texts and composing processes within and across groups, as opposed to the presence or absence of such differences. Thus, we opted for descriptive statistics based on the mean *(M)* and standard deviations *(SD),* along with the effect size expressed with Hedges' *g* and Hedges' *g* adjusted. Using Hedges' *g* is justified as this effect size considers the sample size, applying a correction factor for small sample sizes (Turner & Bernard, 2006). The interpretation of Hedges' *g* is based on Plonsky and Oswald's (2014) benchmarks for the field of L2 research: within-groups (small: .60; medium: 1.0; large: 1.4) and between-groups (small: .40; medium: .70; large: 1.0).

3.3. Results[1]

3.3.1. Fluency: Product measures

Our first research question asked to what extent WCF in the form of a model text affects L2 English children's writing fluency as measured by product indices.

[1] A reviewer suggested including correlational analyses among our fluency measures to check if some of them were redundant. We had already looked into this and found that there was a wide range of correlational values among these measures. Hence our final recommendation in the concluding section.

Tables 1 and 2 provide the descriptive statistics for the fluency product measures within and between groups, respectively.

Table 1 Fluency: Within-group product measures

	Feedback group: *M* (*SD*)				Self-editing group, *M* (*SD*)			
	Initial text	Rewritten text	Mean difference	*g*	Initial text	Rewritten text	Mean difference	*g*
Word count	100.3 (84.63)	104.1 (91.99)	-3.8	-0.04	46.13 (42.29)	40.75 (16.71)	5.38	0.24
Characters/minute	20.66 (14.70)	27.87 (15.29)	-7.21	0.46	16.82 (8.33)	37.92 (13.80)	-21.1	1.75
Words/minute	4.2 (3.17)	5.52 (3.18)	-1.32	0.40	3.22 (1.61)	7.14 (2.96)	-3.92	1.56
Minutes per 100 characters	2.69 (1.78)	2.07 (0.87)	0.62	0.42	2.37 (0.80)	1.37 (0.34)	1.0	1.54
Minutes per 20 words	2.77 (1.97)	2.13 (0.89)	0.64	0.40	2.52 (0.97)	1.5 (0.49)	1.02	1.25

As can be seen in Table 1, the feedback group improved their fluency in all the five product indices from their initial to their rewritten texts, with small effect sizes except for the trivial one in word count. The self-editing group decreased their word count from their initial to their rewritten texts with a small effect size, while the results of the four remaining measures improved, reflecting medium (minutes per 20 words) and large effect sizes (characters and words per minute and minutes per 100 characters).

Table 2 Fluency: Between-group product measures

	Group *M* (*SD*)		Mean difference	$g_{adjusted}$
	Feedback: rewritten text	Self-editing: rewritten text		
Word count	104.1 (91.99)	40.75 (16.71)	63.35	0.08
Characters/minute	27.87 (15.29)	37.92 (13.80)	-10.05	-0.95
Words/minute	5.52 (3.18)	7.14 (2.96)	-1.62	-0.86
Minutes per 100 characters	2.07 (0.87)	1.37 (0.34)	0.7	0.75
Minutes per 20 words	2.13 (0.89)	1.5 (0.49)	0.63	0.66

Between groups (see Table 2), the largest improvements in four product measures were obtained by the self-editing group, whith effect sizes ranging from small (minutes per 20 words) to medium (characters and words per minute and minutes per 100 characters). The only effect size in favor of the feedback group – for word count – yielded a trivial magnitude.

3.3.2. Fluency: Process measures

Our second research question asked to what extent WCF in the form of a model text affects L2 English children's writing fluency as measured by process indices. The descriptive statistics for these measures within and between groups are included in Tables 3 and 4, respectively.

Table 3 Fluency: Within-group process measures

	Feedback group: *M* (*SD*)				Self-editing group: *M* (*SD*)			
	Initial text	Rewritten text	Mean difference	*g*	Initial text	Rewritten text	Mean difference	*g*
Linear fluency I (character-based)	30.81 (16.84)	40.44 (17.51)	-9.63	-0.54	26 (10.82)	47.17 (18.17)	-21.17	-1.34
Linear fluency II (word-based)	5.81 (3.71)	7.03 (3.46)	-1.22	-0.33	4.51 (2.02)	8.48 (3.50)	-3.97	-1.31
Characters/P-Burst	6.87 (3.35)	9.81 (4.76)	-2.94	-0.68	5.76 (2.24)	8.58 (4.25)	-2.82	-0.78
Words/P-Burst	1.3 (0.78)	1.74 (1.06)	-0.44	-0.45	1.01 (0.40)	1.52 (0.75)	-0.51	-0.80
Product/process ratio	.64 (.17)	.68 (.15)	-0.04	-0.20	.65 (.12)	.81 (.05)	-0.16	-1.61

As observed in Table 3, from their initial to their rewritten texts, both groups increased their fluency in all the five process measures. The feedback group displayed a trivial effect size in the product/process ratio and small ones in the four other process indices. In the self-editing group, small effect sizes were found for characters and words per P-burst, the effect sizes in linear fluency I and II yielded medium magnitudes and the results of the product/process ratio showed a large effect size.

Table 4 Fluency: Between-group process measures

	Group *M* (*SD*)		Mean difference	$g_{adjusted}$
	Feedback, rewritten text	Self-editing, rewritten text		
Linear fluency I (character-based)	40.44 (17.51)	47.17 (18.17)	-6.73	-0.68
Linear fluency II (word-based)	7.03 (3.46)	8.48 (3.50)	-1.45	-0.80
Characters/P-Burst	9.81 (4.76)	8.58 (4.25)	1.23	-0.10
Words/P-Burst	1.74 (1.06)	1.52 (0.75)	0.22	-0.21
Product/process ratio	.68 (.15)	.81 (.05)	-0.13	-1.08

Between groups (see Table 4), the self-editing group displayed the largest gains in three process indices: linear fluency I and II and product/process ratio, whose effect sizes ranged from small to medium and large, respectively. The effect sizes in favor of the feedback group were trivial for characters per P-burst and small for words per P-burst.

4. Discussion

Given Truscott's arguments regarding the detrimental and non-detrimental effects (Truscott, 1996 and 2004, respectively) of having access to direct and indirect WCF on fluency in L2 writing, together with the scant research on young learners' digital writing processes, our study intended to shed light on whether processing WCF has an effect on young EFL learners' fluency behavior in digital writing. To this end, we looked at fluency using a multidimensional approach – including product and process perspectives – in the texts typed by L2 children writers, with and without access to WCF shaped as models.

Concerning the product measures, the feedback group increased their total word count after rewriting their texts with the help of models as opposed to the decrease in total words written by the self-editing group. Thus, our data do not seem to support Truscott's (1996) claim that providing WCF results in writing less text, although it should be remembered that Truscott referred to direct and indirect WCF and that the differences here were minimal in raw numbers and in the effect size both within and between groups (see Tables 1 and 2). The results of the self-editing group differ from Ekanayaka and Ellis (2020), whose two direct semi-focused feedback groups increased their text length in the writing of a new text of an identical type to the first one; however, as opposed to our study, the control group increased it too.

The trend of our feedback group's results does not fully coincide with two pen-and-paper studies which used models as WCF. Contrary to our own feedback group and to her collaborative group, there was a decrease in the three quantity fluency measures used by Lázaro-Ibarrola (2021) in young learners' timed individual writing (word count among them), with an absence of any statistically significant results. The rewritten texts of the two proficiency levels in Cánovas Guirao's (2018) non-model-instruction group had similar word counts in untimed collaborative writing after processing models, as opposed to the increase in the model-instruction group, whose texts were always longer than those of the other group. However, the results of our model group coincide with those of Roothooft et al. (2022) in the slight increase of fluency in the texts rewritten by the group who processed models – as opposed to their direct WCF group, whose rewritten texts displayed a slight reduction. Certainly, future research with different age groups should aim to contribute empirical evidence to clarify the effects of each WCF technique and self-editing on writing fluency measured as frequency counts (text length among them). Moreover, a relevant moderating variable of the effect of WCF as models on word count could be not just proficiency level, but including explicit, teacher/researcher-guided instructions about how to use models before students process the WCF and rewrite their texts.

Regarding the remaining four product indices, both groups in our study improved their fluency by increasing their characters and words per minute and by reducing the mean of minutes used to type every 100 characters and 20 words. Our results concur with Sánchez's (2019) self-editing participants in the low-proficiency group (whose level resembles our own participants'), given that they significantly increased their words and syllables per minute in comparison to the direct and indirect WCF groups. Future empirical research in digital writing is warranted to confirm or nuance the aforementioned attested gains in such fluency measures, as mediated by the processing of different WFC techniques and self-editing. Likewise, similar to our model group, the mean of minutes per

100 words from their first to their new fifth assignment was significantly lower in all Chandler's (2003) direct and indirect feedback adult groups in her two studies. In our own one, the largest improvements between groups for both minutes per 100 characters and 20 words were revealed in the self-editing group, with medium and small effect sizes, respectively (see Table 2). The results of the model group, though less beneficial than those of the self-editing group, seem to concur with Truscott's (2004) affirmation that learners improve their speed fluency after processing direct and indirect WCF. Thus, future digital writing studies should provide further empirical evidence about both Truscott's (1996, 2004) contentions and analyze the patterns of the length and speed aspects of fluency that emerge from self-correction versus the processing of different WCF techniques (models, direct and indirect feedback): namely, isomorphic or mixed.

The results of the process measures offer a more mixed picture than that of the product indices (see Tables 3 and 4). The feedback group augmented their fluency in all the five measures, but the self-editing group attained higher mean values than the feedback group for linear fluency I, linear fluency II (characters and words per minute, respectively) and product/process ratio, with effect sizes ranging from small to large. Although the feedback group's mean values obtained in the rewriting session for characters and words per P-burst were higher than those of the self-editing group, the respective trivial and small effect sizes diminish the impact of such results. Again, the descriptive tendencies shown in the present study need to be empirically tested in future digital writing research about the effects of models and other WCF techniques plus self-editing on different fluency process measures.

Our results seem to confirm the plausibility of drawing on Skehan's (2009) application of Levelt's model (1989) of L1 speech production to second/foreign language CAF oral performance in order to study the relationship between models and writing fluency (see section 2.3). Processing model texts slightly increased the feedback group's fluency in all the ten measures. Remembering the lexical alternatives from the comparison stage arguably contributed to easing the pressure on the feedback group's translation process in the rewriting stage, with probable positive effects on their execution too. Besides, similar to the case of oral narratives (Skehan, 2009), it could be argued that the structured nature of the writing task, as reflected in the six-picture sequential prompt available in the rewriting stage, aided to benefit the feedback group's fluency too. However, overall, self-edition turned out to be more efficient for the improvement of fluency.

The consideration of the rewritten phase as an exact task repetition stage (Bygate, 2001) could be useful to jointly explain: 1) the increase of the feedback group's fluency in their rewritten texts and their composing processes and 2) the generally higher fluency improvements of the self-editing group. Certainly, the actual task type and content plus the procedure for rewriting the text were the same as in the initial writing session. From that angle, our results are in line with the revealed tendency that task repetition in the oral and writing modalities appears to impact fluency positively in the repeated task (see Sánchez, 2019, for a review), with greater benefits for the children who did not process the models.

Following Kellogg's (1996) model of writing, the task "rehearsal" of the initial writing and the review of their texts in the second stage allegedly supplied the children with sufficient support or background knowledge, even for the self-editing group (despite their lack of access to the model). These two aspects, coupled with the availability of the six-picture sequential stimulus in the rewriting (task repetition) stage, possibly eased the pressure in the participants' central executive and visual-spatial sketchpad, since they potentially remembered what and in which order they had to write about, and so they did not need to plan at length. In the case of the self-editing group, the absence of processing the model text might have reduced their inclination to pay focused attention to or remember specific linguistic forms in the rewriting stage. Arguably, this factor could have freed up resources in their central executive and phonological loop, thus rendering the linguistic encoding of their preverbal thoughts during the translation process easier and faster. This seems to be supported by their higher reduction of revisions compared to the feedback group (see Table 4 for the results of the product/process ratio measure).

Nevertheless, although the group who did not process the models showed more fluency than the feedback group in seven of the ten measures, it should be considered that the overall higher fluency of the children who self-edited their texts does not necessarily entail more (significant) efficiency in terms of accuracy (see, for instance, Sánchez, 2019 and Lázaro-Ibarrola, 2021). Despite being relevant in any WCF study, this aspect, together with the analysis of complexity, was beyond the scope of this article.

5. Conclusions

The present study was prompted by Truscott's (1996) claim that the provision of direct and indirect WCF involves writing less text, at a higher speed (Truscott, 2004), due to the students' desire to avoid committing mistakes. Specifically, it investigated the effects of the provision of models as WCF and self-editing on the fluency of young L2 learners (measured with product and process indices), an age group whose digital writing processes had not been previously studied.

The results of this study suggest that using model texts does not fully align with the effects produced by direct and indirect WCF on writing fluency according to Truscott's (1996, 2004) claims, given that our model group increased both their word count (though slightly) and speed fluency in the texts they rewrote following the processing of the models. They also augmented their fluency in the remaining eight measures. Between groups, the children who self-edited their texts experienced higher improvements in speed fluency and in six more measures out of the ten indices.

Our study seems to support the theoretical feasibility of applying Skehan's (2009) psycholinguistic explanation about the effects of task design features on processing demands and oral L2 CAF performance to written language. From a pedagogical perspective, an implication of this study is that opting for exact task repetition preceded by processing models as WCF would not be harmful to increase writing fluency. Nevertheless, if teachers want to boost their learners' fluency in rewritten texts, self-correction in general would seem to be the most efficient option.

Several important limitations of this study, which open some lines for further research, should be acknowledged. First, the small size of our sample restricts the generalizability of our findings. Second, we used just one data source and so it would be useful to triangulate the data from keystroke logging programs with the participants' comments on their internal composing processes as elicited by stimulated recalls. Third, we acknowledge that the analysis of accuracy in the initial and final texts before and after processing different techniques of WCF and self-correction would provide valuable information to understand the nature of the relationship between accuracy and fluency in digital texts as mediated by the presence or absence of WCF. Fourth, the analysis of complexity in the same aforesaid terms would complement our knowledge about 1) Truscott's (2004) affirmation of an inverse effect between complexity and speed fluency in texts composed after processing WCF, and 2) any other possible trade-off effects in the full CAF triadic dimension of L2 performance. Importantly, any future research should attempt to control for participants' typing skills, which might mediate the effects of WCF on fluency (especially speed fluency, as a reviewer noted).

Finally, our results lead us to advocate a multidimensional approach (Van Waes & Leijten, 2015) to study writing fluency in a deep and comprehensive manner. Therefore, we would like to highlight the relevance of conducting a robust psychometric study which discriminates the magnitude and role of different product and process fluency measures, for different age groups, with a view to devising composite variables which would avoid multicollinearity and facilitate comparability of findings. In hindsight, we cannot guarantee that multicollinearity did not affect our own results. Overall, given the under-explored domain of young children's digital writing processes in general and of fluency in particular, either alone or as mediated by WCF, we believe that our study represents a worthy contribution, in terms of uncovering empirical questions worth addressing in future studies.

Acknowledgements

This article is part of the I+D+i Grants PID2019-104353GB-I00, funded by MCIN/AEI/10.13039/501100011033, and 20832/PI/18, funded by Fundación Séneca de la Región de Murcia, Spain. We would also like to express our deep gratefulness to the school which allowed us to carry out our study and to the children themselves. Finally, we sincerely thank the two anonymous reviewers and Dr. Rosa M. Manchón for their valuable comments.

References

Abdel Latif, M. M. (2009). Towards a new process-based indicator for measuring writing fluency: Evidence from L2 writers' think-aloud protocols. *Canadian Modern Language Review, 65*(4), 531-558. https://doi.org/10.3138/cmlr.65.4.531

Abdel Latif, M. M. (2013). What do we mean by writing fluency and how can it be validly measured? *Applied Linguistics*, *34*, 99-105. https://doi.org/10.1093/applin/ams073

Baddeley, A. D. (1986). *Working memory*. Oxford University Press.

Barkaoui, K. (2007). Revision in second language writing: What teachers need to know. *TESL Canada Journal, 25*(1), 81-92. https://doi.org/10.18806/tesl.v25i1.109

Barkaoui, K. (2019). What can L2 writers' pausing behavior tell us about their L2 writing processes? *Studies in Second Language Acquisition*, *41*(3), 529-554. https://doi.org/10.1017/s027226311900010x

Breuer, E. O. (2019). Fluency in L1 and FL writing: An analysis of planning, essay writing and final revision. In E. Lindgren & K. P. H. Sullivan (Eds.), *Observing writing: Insights from keystroke logging and handwriting* (pp. 191-211). Brill.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks, second language learning, teaching and testing* (pp. 23-48). Longman.

Cánovas Guirao, J. (2018). *The use of written models in the teaching of English in primary* (Doctoral dissertation, University of Murcia). DIGITUM repository. http://hdl.handle.net/10201/55971

Cánovas Guirao, J., Roca de Larios, J., & Coyle, Y. (2015). The use of models as a written feedback technique with young EFL learners. *System, 52*(1), 63-77. https://doi.org/10.1016/j.system.2015.04.002

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing, 12*(3), 267-296. https://doi.org/10.1016/S1060-3743(03)00038-9

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

Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment.* https://rm.coe.int/1680459f97

Coyle, Y., & Roca de Larios, J. (2020). Exploring young learners' engagement with models as a written corrective technique in EFL and CLIL settings. *System, 95*, 1-14. https://doi.org/10.1016/j.system.2020.102374

Coyle, Y., Cánovas Guirao, J., & Roca de Larios, J. (2018). Identifying the trajectories of young EFL learners across multi-stage writing and feedback processing tasks with model texts. *Journal of Second Language Writing, 42,* 25-43. https://doi.org/10.1016/j.jslw.2018.09.002

Ekanayaka, W. I., & Ellis, R. (2020). Does asking learners to revise add to the effect of written corrective feedback on L2 acquisition? *System, 94,* 102341. https://doi.org/10.1016/j.system.2020.102341

García Hernández, J., Roca de Larios, J., & Coyle, Y. (2017). Reformulation as a problem-solving space for young EFL writers: A longitudinal study of language learning strategies. In M. P. García-Mayo (Ed.), *Learning foreign languages in primary school. Research insights* (pp. 193-222). Multilingual Matters. https://doi.org/10.21832/9781783098118-012

García-Mayo, M. D. P., & Labandibar, U. L. (2017). The use of models as written corrective feedback in English as a foreign language (EFL) writing. *Annual Review of Applied Linguistics*, *37*, 110-127. https://doi.org/10.1017/S0267190517000071

Garcés-Manzanera, A. (2021). *An exploratory study of primary school children's writing processes in digital environments: The use of models as written corrective feedback* (Doctoral dissertation, University of Murcia). DIGITUM repository. http://hdl.handle.net/10201/114544

Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, *11*(4), 459-479. https://doi.org/10.1177/1362168807080963

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Lawrence Erlbaum.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, *30*(4), 461-473. https://doi.org/10.1093/applin/amp048

Kang, E. Y. (2020). Using model texts as a form of feedback in L2 writing. *System, 89*, 102196. https://doi.org/10.1016/j.system.2019.102196

Karim, K., & Nassaji, H. (2020). The effects of written corrective feedback: A critical synthesis of past and present research. *Instructed Second Language Acquisition*, *3*(1), 28-52. https://doi.org/10.1558/isla.37949

Kellogg, R. T. (1990). Effectiveness of prewriting strategies as a function of task demands. *American Journal of Psychology, 103*, 327-342.

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

Kim, Y., Choi, B., Yun, H., Kim, B., & Choi, S. (2022). Task repetition, synchronous written corrective feedback and the learning of Korean grammar: A classroom-based study. *Language Teaching Research*, *26*(6), 1106-1132. https://doi.org/10.1177/1362168820912354

Lázaro-Ibarrola, A. (2021). Model texts in collaborative and individual writing among EFL children: Noticing, incorporations, and draft quality. *International Review of Applied Linguistics in Language Teaching*, 000010151520200160. https://doi.org/10.1515/iral-2020-0160

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

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* MIT Press.

Luquin, M., & García-Mayo, M. P. (2021). Exploring the use of models as a written corrective feedback technique among EFL children. *System, 98,* 1-13. https://doi.org/10.1016/j.system.2021.102465

Martínez-Esteban, N., & Roca de Larios, J. (2010). The use of models as a form of written feedback to secondary school pupils of English. *International Journal of English Studies*, *10*(2), 143-170. https://doi.org/10.6018/ijes/2010/2/119241

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

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

Palviainen, A., Kalaja, P., & Mäntylä, K. (2012). Development of L2 writing: Fluency and proficiency. *AFinLA-e: Soveltavan kielitieteen tutkimuksia, 4,* 47-59.

Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, *64*(4), 878-912. https://doi.org/10.1111/lang.12079

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

Roothooft, H., Lázaro-Ibarrola, A., & Bulté, B. (2022). Task repetition and corrective feedback via models and direct corrections among young EFL writers: Draft quality and task motivation. *Language Teaching Research*, *0*(0). https://doi.org/10.1177/13621688221082041

Sánchez, A. (2019). *External task-repetition: The role of modality, written corrective feedback and proficiency. A comparative study* (Doctoral dissertation, University of Murcia). DIGITUM repository. http://hdl.handle.net/10201/65841

Schoonen, R., Snellings, P., Stevenson, M., & Van Gelderen, A. (2009). Towards a blueprint of the foreign language writer: The linguistic and cognitive demands of foreign language writing. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 77-101). Multilingual Matters.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge University Press.

Semke, H. D. (1984). Effects of the red pen. *Foreign Language Annals, 17,* 195-202. https://doi.org/10.1111/j.1944-9720.1984.tb01727.x

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, *30*(4), 510-532. https://doi.org/10.1093/applin/amp047

Spelman Miller, K., Lindgren, E., & Sullivan, K. P. H. (2008). The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging. *TESOL Quarterly*, *42*, 433-453. https://doi.org/10.1002/j.1545-7249.2008.tb00140.x

Stevenson, M. M. C. (2005). *Reading and writing in a foreign language: A comparison of conceptual and linguistic processes in Dutch and English* (Doctoral dissertation, University of Amsterdam). UvA-DARE (Digital Academic Repository). https://hdl.handle.net/11245/1.310727

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidhofer (Eds.), *For H. G. Widdowson: Principles and practice in the study of language. A Festschrift on the occasion of his 60th birthday* (pp. 125-144). Oxford University Press.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, *46*(2), 327-369. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x

Truscott, J. (2004). Evidence and conjecture on the effects of correction: A response to Chandler. *Journal of Second Language Writing, 13,* 337-343. https://doi.org/10.1016/j.jslw.2004.05.002

Turner, H. M. I., & Bernard, R. M. (2006). Calculating and synthesizing effect sizes. *Contemporary issues in communication science and disorders*, *33*, 42-55. https://doi.org/10.1044/cicsd_33_S_42

Van Gelderen, A., & Oostdam, R. (2002). Improving linguistic fluency for writing: Effects of explicitness and focus of instruction. *L1-Educational Studies in Language and Literature, 2,* 239-270. https://doi.org/10.1023/A:1021304027877

Van Waes, L., & Leijten, M. (2015). Fluency in writing: A multidimensional perspective on writing fluency applied to L1 and L2. *Computers and Composition: An International Journal, 38,* 79-95. https://doi.org/10.1016/j.compcom.2015.09.012

Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity.* University of Hawai'i Press.