This item is the archived peer-reviewed author-version of:

Don't wait until it is too late : the effect of timing of automated feedback on revision in ESL writing

Reference:

Conijn Maria Anna, van Zaanen Menno, Van Waes Luuk.- Don't wait until it is too late : the effect of timing of automated feedback on revision in ESL writing
Transforming Learning with Meaningful Technologies : 14th European Conference on Technology Enhanced Learning, EC-TEL 2019, Delft, The Netherlands,
September 16–19, 2019, Proceedings / Scheffel, M.; Broisin, J.; Pammer-Schindler, V.; Ioannou, A.; Schneider, J. - ISBN 978-3-030-29735-0 - Cham, Springer, 2019, p.
577-581

Full text (Publisher's DOI): https://doi.org/10.1007/978-3-030-29736-7_43

Don't wait until it is too late: The effect of timing of automated feedback on revision in ESL writing

Rianne Conijn^{1,2}[0000–0002–6316–4892], Menno van Zaanen¹[0000–0003–1841–2444],
and Luuk van Waes²[0000–0002–3642–9533]

¹ Department of Cognitive Science and Artificial Intelligence, Tilburg University,
the Netherlands {[m.a.conijn](mailto:m.a.conijn@uvt.nl),[mvzaanen](mailto:mvzaanen@uvt.nl)}@uvt.nl

² Department of Management, University of Antwerp, Belgium
luuk.vanwaes@uantwerpen.be

Abstract. Automated writing evaluation tools have been shown to improve writing quality. However, the impact of automated feedback, and especially the timing of the feedback, on students' writing process is still unknown. Hence, we analyzed how feedback timing influences the revision process. Three experimental conditions were implemented into the writing tool CyWrite: no feedback, immediate feedback during the drafting and revision stage, and feedback during the revision stage only. Keystroke data were collected from 60 ESL students while conducting a source-based argumentative writing task. The revisions made during the writing process and the students' satisfaction with the system were analyzed. The results showed little differences in the amount, size, and duration of revisions between the three conditions. However, students reported they felt more disrupted when feedback was provided during the full writing process rather than in the revision stage only.

Keywords: Feedback · Writing analytics · Keystroke analysis · Revision
· Automated writing evaluation

1 Introduction

Receiving timely and personalized feedback on writing is very useful for (learner) writers, but providing it is a time-intensive task for teachers. Therefore, a wide variety of writing tools have been developed to assist teachers by providing automated grading and feedback on students' writing. They can be classified as follows: automated essay scoring (AES), automated writing evaluation (AWE), and intelligent tutoring systems (ITS) [1]. AES are mostly used for summative feedback [4], AWE for formative feedback [3], and ITS include instructional elements and interactivity next to the feedback [10]. Several researchers started to evaluate these systems both on the accuracy of the automated feedback provided (e.g., [6]) as well as on the effectiveness of the tools used in classroom settings (e.g., [13]). Overall, the automated feedback components have fairly high accuracies [1]. In addition, most of these tools have been shown to improve writing quality and enhance student motivation and autonomy [3].

Most of these evaluations, however, evaluate the impact on the writing product, rather than the impact on the writing process [3, 14]. Hence, we cannot determine *how* the tool is used, for example, whether the automated feedback leads to more revisions. Some studies did indicate an effect of feedback on the revision of errors, but these are mostly descriptive case studies, e.g., [5]. In addition, it has been argued that for (automated) feedback to be effective it needs to be timely and frequently [1, 3, 7]. However, little is known about the impact of the timing of automated feedback on the writing process. Therefore, in the current study we aim to identify how the timing of automated feedback during English as a Second Language (ESL) writing impacts the frequency, level, and duration of revisions.

2 Method

In this study, participants were asked to write an argumentative text on global warming of 250–350 words, by using two short sources of text (250 words each). The participants were allowed 20 minutes to read the sources and write the text (drafting stage). Thereafter, the participants received unlimited time to revise their text (revision stage). In total, 60 ESL undergraduate students participated in this study. The writing task was conducted using the web-based AWE tool CyWrite [12]. CyWrite uses both statistical and rule-based natural language processing to provide formative feedback on spelling, grammar, style, and discourse patterns [2, 6]. This feedback is generated during the writing process, by underlining errors and providing feedback in the margin. A screenshot of CyWrite with two examples of feedback (spelling in red, grammar in green) is shown in Figure 1. The participants were randomly divided into three conditions (each $N = 20$) regarding the timing of the feedback: 1) no feedback; 2) feedback during the revision stage only; 3) immediate feedback during the full writing process (both drafting and revision stages).

During the writing task, keystroke data were collected using CyWrite. After the writing task, a questionnaire collected students' demographics, writing style [8], and satisfaction with the feedback [?]. For the current study, only the keystroke data and satisfaction data were analyzed. The number of revisions, size of the revision, and duration of the revisions were extracted using R. The size of the revision was calculated by the number of word-level (as opposed to sub-word level) revisions [11]. ANOVAs were used to analyze the differences in the frequency and size of the revisions and the duration of the revision stage between the three conditions. In addition, ANOVAs were used to determine the differences in satisfaction between the two feedback conditions.

3 Results

On average, the participants typed 1999 ($SD = 517$) characters, resulting in a final text of 258 ($SD = 64$) words. All students, except for six students in the no feedback condition, revised during the revision stage. The descriptive

Fig. 1. Screenshot of CyWrite with formative feedback.

Table 1. Descriptive statistics for the three conditions and the full dataset ($N = 60$).

Feedback	No feedback	Revision stage	Immediate	All data
	Mean (<i>SD</i>)	Mean (<i>SD</i>)	Mean (<i>SD</i>)	Mean (<i>SD</i>)
#Revisions drafting stage	88.5 (53.9)	102 (50.0)	92.8 (40.8)	94.5 (48.0)
#Revisions revision stage	17.0 (19.7)	29.7 (26.9)	28.2 (21.2)	25.0 (23.1)
#Word-level revisions drafting	29.9 (15.1)	32.6 (14.4)	27.8 (12.6)	30.1 (14.0)
#Word-level revisions revision	7.90 (9.89)	7.45 (8.80)	11.1 (8.61)	8.80 (9.11)
Duration revision stage (sec)	247 (233)	347 (233)	403 (284)	333 (255)

statistics for the revision behavior per condition are shown in Table 1. Little differences were found in the number of revisions, especially in the drafting stage. The number of revisions in the revision stage was higher when feedback was provided, compared to no feedback, but this difference was not significant ($F(2,57) = 1.86$, $p = .17$, $\eta^2 = .06$). In addition, the revision stage was longer when feedback was provided in the revision stage, compared to no feedback, and even longer when feedback was immediately provided. However, this difference was also not significant ($F(2,57) = 1.99$, $p = .15$, $\eta^2 = .07$). No differences were found between satisfaction with the feedback in the two feedback conditions ($F(1,38) = 0.57$, $p = .46$, $\eta^2 = .01$). Interestingly, when immediate feedback was provided, students felt more disrupted ($F(1,38) = 25.25$, $p < .001$, $\eta^2 = .40$).

4 Discussion and Conclusion

In the current study we aimed to identify the effect of timing of automated feedback on how ESL students revise. A trend was found that students revised more when automated feedback was provided, compared to no feedback. In addition, students took more time to revise when immediate feedback was provided during the full writing process, compared to when students received no feedback or only feedback during the revision stage. However, these differences were not found significant. In addition, feedback during the full writing process did not result in higher satisfaction, but was considered significantly more disruptive compared to feedback during the revision stage only. Thus, although timely feedback has been argued to be most useful [1], this is not clearly reflected in the revision patterns nor the users' satisfaction. These results can be explained by the high

variance in the revisions made between the students. This is in line with previous work that also indicated that nature and size of the effect of automated feedback on revision differs across ESL students [5]. Hence, larger sample sizes or more insight into individual differences might be necessary to determine the effect of (the timing of) feedback on revisions.

This study showed a first step into analyzing the effect of automated writing feedback on the writing process, and specifically on how ESL students revise. The current results are inherently bound by the content and format of the feedback as provided in the CyWrite system. Future work should investigate the effect of the content or format of the feedback on the revision process. Current writing tools are often criticized based on the low-level of the feedback they provide. Future work should investigate whether this also leads to low-level revisions, in terms of, for example, depth, immediacy, or recursiveness of the revision, see Lindgren and Sullivan's revision taxonomy [9].

References

1. Allen, L.K., Jacovina, M.E., McNamara, D.S.: Computer-based writing instruction. *Handbook of writing research* pp. 316–329 (2015)
2. Chukharev-Hudilainen, E., Saricaoglu, A.: Causal discourse analyzer: Improving automated feedback on academic esl writing. *Computer Assisted Language Learning* **29**(3), 494–516 (2016)
3. Cotos, E.: Automated writing analysis for writing pedagogy: From healthy tension to tangible prospects. *Writing and Pedagogy* **6**, 1 (2015)
4. Dikli, S.: An overview of automated scoring of essays. *The Journal of Technology, Learning and Assessment* **5**(1) (2006)
5. El Ebyary, K., Windeatt, S.: The impact of computer-based feedback on students written work. *International Journal of English Studies* **10**(2), 121–142 (2010)
6. Feng, H.H., Saricaoglu, A., Chukharev-Hudilainen, E.: Automated error detection for developing grammar proficiency of esl learners. *calico journal* **33**(1), 49 (2016)
7. Ferguson, P.: Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education* **36**(1), 51–62 (2011)
8. Kieft, M., Rijlaarsdam, G., van den Bergh, H.: An aptitude-treatment interaction approach to writing-to-learn. *Learning and Instruction* **18**(4), 379–390 (2008)
9. Lindgren, E., Sullivan, K.: *Analysing online revision*, pp. 157–188. Elsevier (2006)
10. Ma, W., Adesope, O., Nesbit, J., Liu, Q.: Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology* **106**(4), 901 (2014)
11. Monahan, B.D.: Revision strategies of basic and competent writers as they write for different audiences. *Research in the Teaching of English* pp. 288–304 (1984)
12. Ranalli, J.R., Feng, H.H., Chukharev-Khudilaynen, E.: The affordances of process-tracing technologies for supporting l2 writing instruction. *Language Learning & Technology* **23**(2) (2018)
13. Roscoe, R.D., Allen, L.K., Weston, J.L., Crossley, S.A., McNamara, D.S.: The writing pal intelligent tutoring system: Usability testing and development. *Computers and Composition* **34**, 39–59 (2014)
14. Stevenson, M., Phakiti, A.: The effects of computer-generated feedback on the quality of writing. *Assessing Writing* **19**, 51–65 (2014)