
ceived ICF in relation to the Grammar and Spelling features.

5.2.3 Student Perception of the Feedbacks

After the study, students in both classes were asked to separately identify their level of agreement on that the feedback is helpful on each of seven essay features. The five-point likert scale rate was used to rate the level of agreement. In addition to the questionnaire, informal user feedback was collected from students in both classes.

In the DCF class, more than 73% participants agreed or strongly agreed that the feedback was helpful on *Grammar, Spelling, Conclusion, Supporting Ideas, Coherence and organization*; 62% participants agreed or strongly agreed that the feedbacks were helpful on *sentence diversity*.

Most of the students interviewed in the DCF class said that the DCF was helpful in the sense that it pointed out their errors in writing, so that student knew what errors they had made. However, when asked if the feedback was useful to improve their writing skills, doubts were expressed: *Sometimes, I don't understand teacher comments, so just quickly accept the changes suggested by the teacher. I might make the same mistakes in the next composition.*

In the ICF class, more than 65% participants agreed or strongly agreed that the feedback was helpful on *Conclusion, Supporting Ideas, Coherence and Organization*. Forty-eight percent of participants agreed or strongly agreed that the feedback was helpful on *grammar, spelling and sentence diversity*. Some students interviewed said that they liked the ways to explore solutions by themselves. However, other students hoped that the feedback could be more specific. Moreover, students complained that the feedbacks were too numerous if four to seven aspects of their essays were required to double check.

Overall, most of students in the DCF class agreed with the usefulness of the feedback on seven aspects of the essay, while those students in the ICF class found that the feedback was particularly useful in supporting ideas, organization, coherence and conclusion. Less positive feedback on *Grammar, Spelling and Sentence Diversity* in the ICF class were received since they were too general. This result is in line with Miceli's study [15], where students felt that ICF was useful in encouraging them to reflect on aspects of their writing and to develop improvements in the content, while DCF to be more helpful when revising syntax and vocabulary.

5.2.4 Observations of The Computer Generated Revision History

The system keeps a revision history of each document. Based on this information, we can analyze how much text changes were made in terms of number of revisions and text edit, when student received the feedback. Text edit refers to create, modified or remove text action. Table 6

shows that students receiving ICF made more text changes than those students getting DCF regarding to the number of edited words (52 in ICF class, 36 in DCF class) and the number of revisions (108 in ICF class, 79 in DCF class). Independent sample t-test results revealed that students in ICF class made significantly more changes than those students in DCF class, $t(108)=5.798, p < 0.001$ in the number of edited words; $t(108)=10.600, p < 0.001$ in the number of revisions. These results indicated that students in the ICF class spent more time on revising the essay than those students in the DCF class did. The main reason for this is that the ICF encourages students to be more active in their use of feedback [12].

6 CONCLUSION AND FUTURE WORK

The most significant findings of this study indicated that the system-generated ICF was useful to improve the quality of writing, particularly in *the structure, organization, supporting ideas, coherence, and conclusion* aspects of the essay in the short term. In addition, the ICF encouraged students to spend more time on performing self-correction than the DCF provided by teachers. These finds were consisted with previous study results, which reported positive impacts of ICF on the ability of students to edit their own composition and to improve levels of accuracy in writing [10], and the effectiveness of the combined both form and content focused feedback in improving the writing development [4], [8]. Although findings on the effectiveness of different feedback types have been conflicting, largely due to the widely varying student populations, types of writing and feedback practices examined [9], this study result implied that the system could be useful for Chinese English-major students with advanced metalinguistic knowledge, particularly in content development. We also observed that some incorrect ICF were generated. But, the incorrect feedback might be particularly beneficial if they promote noticing. The system required the students to consider the weakness of the essay and evaluate them in context to determine whether they are correct.

This study has some limitations. The system-generated feedback could be too general for correcting errors in some aspects of essay, such as Grammar and Spelling. Moreover, some essays triggered feedback on many features, which generated a daunting number of suggestions. This quantity of feedback seemed to undermine our scaffolding goal by targeting too many essay elements at once. Finally, this study did not examine the impact of the tool in non-English major students' writing. Non-English major students do not have good metalinguistic knowledge, so the impact of the feedback might be different.

In the future work, we will focus on how to combine both DCF and ICF together to support different aspects of

TABLE 6: TEXT CHANGES MADE IN THE REVISION

Class	N	Average Number of Edited Words	Std. Dev. Of Edited Words	Average Number Of Revisions	Std. Dev. Of Revisions
ICF	56	52	20.19	108	21.69
DCF	54	36	12.56	79	14.54

students' writing. Particularly, we will examine different English parsers, such as Stanford Parser, which it might give more specific feedback on spelling and grammar. We also will study how much amount of feedback is sufficient. Zacharias [57] found that if students had too much feedback they would feel discouraged and were less likely to be motivated to use it for revision. This could be limited based on the severity and/or number of negative features detected in the text. Lastly, the future work will investigate how to effectively use this tool for teaching and encouraging independent editing skills. Ferris [58] and Mu [59] have provided descriptions of procedures for helping students learn to self-edit, and computer-assisted feedback could be included in such procedures.

ACKNOWLEDGEMENT

This work is supported by Chongqing Social Science Planning Fund Program under grant No. 2014BS123, Fundamental Research Funds for the Central Universities under grant No. SWU114005, XDJK2017C024, XDJK2017D064, CQU903005203326, and National Natural Science Foundation of China (61502397), and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry, and the Research Program Funds of Faculty of Education, at Southwest University (Learning Technology Assisted Writing Research Project) and the Research Program Funds of the Collaborative Innovation Center of Assessment toward Basic Education Quality at Beijing Normal University (2016-06-016-BZK01).

REFERENCES:

[1] N. Bureau of Statistics of China, *China Statistical YearBook*. China Statistics Press, 2013.

[2] L. Brannon and C. H. Knoblauch, "On students' rights to their own texts: A model of teacher response," *Coll. Compos. Commun.*, vol. 33, no. 2, pp. 157–166, 1982.

[3] I. Leki, "The preferences of ESL students for error correction in college-level writing classes," *Foreign Lang. Ann.*, vol. 24, pp. 203–218, 1991.

[4] A. K. Fathman and E. Whalley, "Teacher Response to Student Writing- Focus on Form versus Content," in *Second Language Writing: Research Insights for the Classroom*, 1990, pp. 178–190.

[5] D. Ferris, *Response to student writing: Implications for second language students*. Mahwah, NJ : Lawrence Erlbaum Associates, 2003.

[6] I. Lee, "Error correction in L2 secondary writing classrooms: The case of Hong Kong," *J. Second Lang. Writ.*, vol. 13, pp. 285–312, 2004.

[7] C. G. van Beuningen, N. H. de Jong, and F. Kuiken, "The Effect of Direct and Indirect Corrective Feedback on L2 Learners' Written Accuracy," *ITL Int. J. Appl. Linguist.*, vol. 156, pp. 279–296, 2008.

[8] T. Ashwell, "Patterns of Teacher Response to Student Writing in a Multiple-Draft Composition

Classroom: Is Content Feedback Followed by Form Feedback the Best Method?," *J. Second Lang. Writ.*, vol. 9, no. 3, pp. 227–257, 2000.

[9] D. R. Ferris, "Does error feedback help student writers? New evidence on the short- and long-term effects of written correction," in *Feedback on Second Language Writing: Contexts and Issues*, K. Hyland and F. Hyland, Eds. New York, 2006, pp. 1–104.

[10] D. Ferris and B. Roberts, "Error feedback in L2 writing classes How explicit does it need to be?," *J. Second Lang. Writ.*, vol. 10, no. 3, pp. 161–184, 2001.

[11] J. Chandler, "The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing," *J. Second Lang. Writ.*, vol. 12, no. 3, pp. 267–296, 2003.

[12] F. Hyland, "Providing effective support : investigating feedback to distance language learners," *Openg Learn.*, vol. 16, no. 3, pp. 233–237, 2001.

[13] D. Little, "Learner autonomy and second/foreign language learning," *Subject Centre for Languages, Linguistics and Area Studies*, 2003. [Online]. Available: <http://www.llas.ac.uk/resources/gpg/1409>.

[14] D. Ferris, *Treatment of error in second language student writing*. Ann Arbor: The University of Michigan Press, 2002.

[15] T. Miceli, "Foreign Language Students' Perceptions of Reflective Approach to Text Correction," *Finders Univ. Lang. Gr. Online Rev.*, vol. 3, no. 1, pp. 25–36, 2006.

[16] Y. Attali and J. Burstein, "Automated Essay Scoring With e-rater V.2.," *J. Technol. Learn. Assess.*, vol. 4, no. 3, 2006.

[17] S. Elliot, "Intellimetric: From here to validity," in *Automated essay scoring: A cross-disciplinary perspective*, Mahwah, NJ: Lawrence Erlbaum Associates, 2003, pp. 71–86.

[18] J. Burstein, M. Chodorow, and C. Leacock, "Automated essay evaluation: The Criterion online writing service," *AI Mag.*, vol. 25, p. 27, 2004.

[19] Richard Haswell, "The complexities of responding to student writing; or, looking for shortcuts via the road of excess," *Across Discip.*, vol. 3, 2006.

[20] S. C. Weigle, "English as a second language writing," in *Handbook of automated essay evaluation*, 2013, pp. 36–53.

[21] K. Yancey, A. Lunsford, J. McDonald, C. Moran, M. Neal, C. Pryor, D. Roen, and C. Selfe, "CCCC position statement on teaching, learning, and assessing writing in digital environments," *Coll. Compos. Commun.*, vol. 55, no. 4, pp. 785–790, 2004.

[22] J. Villalon, P. Kearney, R. A. Calvo, and P. Reimann, "Glosser: Enhanced Feedback for Student Writing Tasks," 2008, pp. 454–458.

[23] M. Liu, R. A. Calvo, and V. Rus, "Automatic Question Generation for Literature Review Writing

- Support,” 2010.
- [24] D. S. McNamara, S. a Crossley, and R. Roscoe, “Natural language processing in an intelligent writing strategy tutoring system.,” *Behav. Res. Methods*, vol. 45, no. 2, pp. 499–515, Jun. 2013.
- [25] M. Liu, R. Calvo, and V. Rus, “Automatic Generation and Ranking of Questions for Critical Review,” *Educ. Technol. Soc.*, vol. 17, no. 2, pp. 333–346, 2014.
- [26] T. K. Landauer, D. S. McNamara, S. Dennis, and W. Kintsch, *Handbook of Latent Semantic Analysis*. Lawrence Erlbaum, 2007.
- [27] L. M. Rudner, “Reducing Errors Due to the Use of Judges.,” *Pract. Assessment, Res. Eval.*, vol. 3, no. 3, 1992.
- [28] S. Dikli, “Automated Essay Scoring,” *Turkish Online J. Distance Educ.*, vol. 7, no. 1, pp. 49–62, 2006.
- [29] M. A. Hearst, “The debate on automated essay grading,” *Intell. Syst. Their Appl.*, vol. 15, no. 5, pp. 22–37, 2000.
- [30] D. Wade-Stein and E. Kintsch, “Summary Street: Interactive computer support for writing,” *Cogn. Instr.*, vol. 22, no. 3, pp. 333–362, 2004.
- [31] M. Shermis and J. Burstein, *Automated Essay Scoring: A Cross-Disciplinary Perspective*. Mahwah NJ: Lawrence Erlbaum Associates, 2003.
- [32] P. F. Ericsson and R. Haswell, “Machine Scoring of Human Essays: Truth and Consequences.” Utah State University Press, 2006.
- [33] G. Gibbs and C. Simpson, “Conditions under which assessment supports students’ learning Learning and Teaching in Higher Education,” *Learn. Teach. High. Educ.*, vol. 1, no. 1, pp. 3–31, 2004.
- [34] R. A. CALVO and R. A. ELLIS, “Students’ Conceptions of Tutor and Automated Feedback in Professional Writing.,” *J. Eng. Educ.*, vol. 99, pp. 427–438, 2010.
- [35] E. C. Thiesmeyer and J. E. Theismeyer, *Editor: A System for Checking Usage, Mechanics, Vocabulary, and Structure*. New York, New York, USA: Modern Language Association, 1990.
- [36] J Anderson, *Mechanically Inclined: Building Grammar, Usage, and Style into Writer’s Workshop*. Stenhouse Publishers, 2005.
- [37] T. J. Beals, “Between Teachers and Computers: Does Text-Checking Software Really Improve Student Writing?,” *English J.*, pp. 67–72, 1998.
- [38] M. A. Britt, P. Wiemer-Hastings, A. A. Larson, and C. A. Perfetti, “Using Intelligent Feedback to Improve Sourcing and Integration in Students’ Essays,” *Int. J. Artif. Intell. Ed.*, vol. 14, no. 3,4, pp. 359–374, 2004.
- [39] T. Kakkonen and E. Sutinen, “EssayAid: towards a semi-automatic system for assessing student texts,” *Int. J. Contin. Eng. Educ. Life-Long Learn.*, vol. 21, no. 2/3, pp. 119–139, 2011.
- [40] A. C. Graesser, D. S. McNamara, M. M. Louwerse and Z. Cai, “Coh-matrix: analysis of text on cohesion and language.,” *Behav. Res. methods, instruments, Comput.*, vol. 36, no. 2, pp. 193–202, May 2004.
- [41] J. W. Pennebaker and M. E. Francis, *Linguistic inquiry and word count (LIWC)*. Mahwah, NJ: Erlbaum, 1999.
- [42] R. M. Rufenacht, P. M. McCarthy, and T. A. Lamkin, “Fairy Tales and ESL Texts: An Analysis of Linguistic Features Using the Gramulator,” in *Proceedings of the Twenty-Fourth International Florida Artificial Intelligence Research Society Conference*, 2011, pp. 287–292.
- [43] S. Crossley and D. McNamara, “Predicting second language writing proficiency : The role of cohesion , readability , and lexical difficulty,” *J. Res. Read.*, vol. 35, pp. 115–135, 2012.
- [44] S. a. Crossley and D. S. McNamara, “Text Coherence and Judgments of Essay Quality: Models of Quality and Coherence,” in *The 33rd Annual Conference of the Cognitive Science Society*, 2011.
- [45] C. Alderson, *Diagnosing foreign language proficiency: The interface between learning and assessment*. London, UK: Continuum, 2005.
- [46] B. North, “Scales for rating language performance: Descriptive models, formulation styles, and presentation formats,” *TOEFL Monogr.*, vol. 24, 2003.
- [47] U. Knoch, “Rating scales for diagnostic assessment of writing: What should they look like and where should the criteria come from?,” *Assess. Writ.*, vol. 16, no. 2, pp. 81–96, 2011.
- [48] I. Lee, “Feedback in Hong Kong secondary writing classrooms: Assessment for learning or assessment of learning?,” *Assess. Writ.*, vol. 12, no. 3, pp. 180–198, 2007.
- [49] S. Crossley and D. McNamara, “Cohesion, coherence, and expert evaluations of writing proficiency,” in *The 32nd Annual Conference of the Cognitive Science Society*, 2010, pp. 984–989.
- [50] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, “The WEKA data mining software,” *ACM SIGKDD Explor.*, vol. 11, pp. 10–18, 2009.
- [51] D. S. . McNamara, A. C. . c Graesser, P. M. . McCarthy, and Z. . Cai, *Automated evaluation of text and discourse with Coh-Matrix*. 2012.
- [52] W. Kintsch and T. van Dijk, “Towards a model of text comprehension and production,” *Psychol. Rev.*, vol. 85, pp. 363–394, 1978.
- [53] J. Lafferty, D. Sleator, and D. Temperley, “Grammatical Trigrams: A Probabilistic Model of Link Grammar,” in *Proceedings of the AAAI Conference on Probabilistic Approaches to Natural Language*, 1992.
- [54] S. a. Crossley and D. S. McNamara, “Understanding expert ratings of essay quality: Coh-Matrix analyses of first and second language writing,” *Int. J. Contin.*

Eng. Educ. Life-Long Learn., vol. 21, no. 2/3, p. 170, 2011.

- [55] L. Jinghong, "The impact of the English Writing Theories on Chinese English Second Language Writing," *Foreign Lang. Teach.*, no. 2, pp. 41–47, 2006.
- [56] J. Hinds, "Quasi-Inductive: Expository Writing in Japanese, Korean, Chinese and Thai," in *Coherence in Writing Research and Pedagogical Perspectives*, 1990.
- [57] N. T. Zacharias, "Teacher and Student Attitudes toward Teacher Feedback," *RELC J.*, vol. 38, no. 1, pp. 38–52, 2007.
- [58] D. R. Ferris, H. Liu, A. Sinha, and M. Senna, "Written corrective feedback for individual L2 writers," *J. Second Lang. Writ.*, vol. 22, no. 3, pp. 307–329, 2013.
- [59] C. Mu, "A Taxonomy of ESL Writing Strategies," *Proc. Redesigning Pedagog. Res. Policy, Pract.*, pp. 1–10, 2005.



Dr. Weiwei Xu is associate professor at College of International Studies, Southwest University in China. She holds a doctorate degree in English from Macquarie University, Australia. She acquired her MA degree in English from Newcastle University, UK. She has been teaching English academic writing for 8 years.



Dr. Li Liu is associate professor at Chongqing University. He is also serving as a Senior Research Fellow of School of Computing at the National University of Singapore. Li received his Ph.D. in Computer Science from the Université Paris-sud XI in 2008. He had served as an associate professor at Lanzhou University in China. His research interests are in pattern recognition,

data analysis, and their applications on human behaviors. He aims to contribute in interdisciplinary research of computer science and human related disciplines. Li has published widely in conferences and journals with more than 50 peer-reviewed publications. Li has been the Principal Investigator of several funded projects from government and industry.



Dr. Ming Liu is Associate Professor at School of Computer and Information Science, Southwest University, China. He received the PhD in Artificial Intelligence in Education at the School of Electrical and Information Engineering, The University of Sydney, Australia in 2012. His main research interests include question generation, learning analytics and intelligent tutoring system. He participated in national and international projects funded by ARC Linkage (Australia), Young and Well CRC, Office of Teaching and Learning, Google and Chinese National Fund. He is an author of over 20 publications papers in prestigious conferences and journals, such as Intelligent Tutoring Systems, IEEE transactions on Learning Technologies, Journal of Educational Technology and Society.

analytics and intelligent tutoring system. He participated in national and international projects funded by ARC Linkage (Australia), Young and Well CRC, Office of Teaching and Learning, Google and Chinese National Fund. He is an author of over 20 publications papers in prestigious conferences and journals, such as Intelligent Tutoring Systems, IEEE transactions on Learning Technologies, Journal of Educational Technology and Society.



Dr. Yi Li is an associate professor at Faculty of Education in the Southwest University in China. She has a master in applied statistics and a PhD in educational measurement by the Purdue University in the USA. Her research interest is application of quantitative methodology

to explore, understand and influence the technology innovation in the field of education. She has participated in the national and international projects funded by National Science Foundation and National Social Science Foundation of China. Yi is author of many research publications in prestigious international conferences and journals.