



# Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction

Jinrong Li<sup>a,\*</sup>, Stephanie Link<sup>b</sup>, Volker Hegelheimer<sup>b</sup>

<sup>a</sup> Department of Writing and Linguistics, Newton Building 2221A, Georgia Southern University, P.O. Box 8026, Statesboro, GA 30460, United States

<sup>b</sup> Department of English, Mailbox 206 Ross Hall, Iowa State University, Ames, IA 50011, United States

## Abstract

The development of language processing technologies and statistical methods has enabled modern automated writing evaluation (AWE) systems to provide feedback on language and content in addition to an automated score. However, concerns have been raised with regard to the instructional and assessment value of AWE in writing classrooms. The findings from a few classroom-based studies concerning the impact of AWE on writing instruction and performance are largely inconclusive. Meanwhile, since research provides favorable evidence for the reliability of AWE corrective feedback, and that writing accuracy is both an important and frustrating issue, it is worthwhile to examine more specifically the impact of AWE corrective feedback on writing accuracy. Therefore, the study used mixed-methods to investigate how Criterion<sup>®</sup> affected writing instruction and performance. Results suggested that Criterion<sup>®</sup> has led to increased revisions, and that the corrective feedback from Criterion<sup>®</sup> helped improve accuracy from a rough to a final draft. The potential benefits were also confirmed by the instructors' interviews. The students' perspectives were mixed, but the extent to which the views vary may depend on the students' language proficiency level and their instructors' use and perspectives of AWE.

© 2014 Elsevier Inc. All rights reserved.

*Keywords:* AWE; ESL writing; Corrective feedback; Mixed-methods research

## Introduction

Since the first automated essay scoring (AES) system in 1966 (Page, 2003, p. 43), increasingly sophisticated language processing technologies and statistical methods have enabled newly developed scoring engines, such as e-rater<sup>®</sup>,<sup>1</sup> Knowledge Analysis Technologies<sup>TM</sup>,<sup>2</sup> and IntelliMetric<sup>®</sup><sup>3</sup> to analyze a wide range of text features at lexical, syntactic, semantic, and discourse levels. Powered by these scoring engines, some automated writing evaluation (AWE) tools, such as Criterion<sup>®</sup> and MY Access!<sup>®</sup>, have been developed to provide formative feedback on

\* Corresponding author. Tel.: +1 515 441 3936.

E-mail address: [jli@georgiasouthern.edu](mailto:jli@georgiasouthern.edu) (J. Li).

<sup>1</sup> <http://www.ets.org/erater/about/>

<sup>2</sup> <http://www.pearsonassessments.com/pai/ai/Products/NextGeneration/ScoringandReporting.htm>

<sup>3</sup> <http://www.vantagelearning.com/products/intellimetric/>

various rhetorical (e.g., organization) and language-related dimensions (e.g., grammar and mechanics) and supplementary resources in addition to automated scores. Although most AES systems have been supported by favorable validity evidence based on the consistency and agreement between the automated system and human raters (Enright & Quinlan, 2010; Keith, 2003; Vantage Learning, 2006; Weigle, 2010), the potential instructional and assessment value of AWE still remains to be examined. Questions have been raised that go beyond the scope of previous psychometric research. For example, researchers believe that high correlation and agreement between AES scores and human ratings is an insufficient condition for the validity of score use (Chung & Baker, 2003, p. 29) and the purpose of such use needs to be taken into account (Chung & Baker, 2003; Weir, 2005). Moreover, critics have expressed concerns over the relevance of the features evaluated by AWE to the true qualities of writing, particularly the social and communicative dimensions (Ericsson, 2006). Recent controversies are increasingly concerned with the use of AWE corrective feedback. The developers and some researchers believe that the automated feedback could help students revise and improve writing while others caution that by diverting students' attention from content development to formal attributes, AWE may lead to negative washback effects on students' writing behavior (CCCC, 2006; Cheville, 2004).

These concerns, together with the lack of adequate evidence of the use of AWE corrective feedback in ESL writing classrooms and its impact on students' writing, prompted the current study. Specifically, the present study explores the role of AWE corrective feedback in an ESL writing curriculum in which one such online writing evaluation service, Criterion<sup>®</sup>, has been integrated into the instructional framework of university ESL writing classrooms. We examine the use and the impact of the automated feedback through a naturalistic classroom-based approach. Responding to a recent call for AWE research initiated by Ware (2011), we center our discussion around the need for clearer understanding of “the context in which the system is used, the content of what is written, and the impact on key stakeholders as part of its integration” (p. 773). Next, we review key classroom-based studies that have contributed to AWE research to date. We then outline our theoretical framework, research questions, methodology, and important findings before we end with a discussion of future directions for the field of L2 writing in general and computer-generated feedback in particular.

## Previous research on AWE

The narrow scope of psychometric research and the lack of studies that examine the use of AWE as it is systematically integrated into writing instruction have opened up a new research agenda, which was taken up by a few studies exploring the effectiveness of AWE on improving students' standardized writing test scores (Attali, 2004; Vantage Learning, 2007), instructors' and students' use of AWE (Chen & Cheng, 2008; Grimes & Warschauer, 2010; Wang, Shang, & Briody, 2012; Warschauer & Grimes, 2008), and their attitudes toward its use in writing courses (Chen & Cheng, 2008; Lai, 2009).

Different measures have also been used to assess more directly the effectiveness of AWE in helping students revise. Although they seemed to have provided some insights into the effects of AWE, the results were far from conclusive. For example, Attali (2004) used the number of submissions as an indicator of students' use of an AWE system and found that 71% of the essays were submitted only once (p. 4). The number of submissions provides a rough estimate of the frequency of students' use of AWE, but any inferences about how students used AWE or benefited from using it based only on the number of submissions would be inaccurate.

To address the issues concerning the process of using AWE, an increasing number of studies have relied on classroom observations and interviews with teachers and students. Chen and Cheng (2008), for instance, examined the effectiveness of MY Access!<sup>®</sup> in assisting writing improvement in three EFL writing classes mostly through interviews with the teachers and the students. By looking into the process, the authors were able to demonstrate a sharp contrast between the three classes in terms of the teacher's use and requirements of using the system and the students' response, and possible connections between the teacher's pedagogical decisions and the students' perception. Meanwhile, the process-oriented approach also allowed the researchers to reveal the complexity in evaluating the effects of AWE. While the study revealed that some students saw the automated feedback as vague and formulaic, others held that it was helpful for identifying and correcting grammatical and mechanical errors, particularly for students with lower English proficiency. Chen and Cheng (2008) cautioned that a quick conclusion about the usefulness of AWE might not be accurate but did not include analysis of writing samples to further the investigation of the issue. The process-oriented approach also unveiled another complexity in evaluating the effects of AWE: Most

Download English Version:

<https://daneshyari.com/en/article/364045>

Download Persian Version:

<https://daneshyari.com/article/364045>

[Daneshyari.com](https://daneshyari.com)