A Comparative Analysis of Formative Feedback Provided by ChatGPT and Human Evaluators in EFL Writing

Saturday, 27 July 2024 15:00 (45 minutes)

Delivering formative feedback on EFL students'writing is an effective method for fostering their writing skills (Graham, Harris, & Hebert, 2011; MacArthur, 2016). Recent advancements in artificial intelligence (AI), such as ChatGPT, could serve as a tool for automated writing assessment, potentially amplifying the volume of feedback students receive and easing the workload for teachers who must frequently provide feedback to sizable classes. This study investigates the feasibility of using generative AI, specifically ChatGPT, as a source of formative feedback for writing instruction, comparing its efficacy with human evaluators. The research examines feedback provided by ChatGPT and human raters on essays written by EFL students who were supposed to achieve a B2 level of proficiency (CEFR). The research analyzed Fifty pieces of human-generated formative feedback and fifty pieces of AI-generated formative feedback for the same essays. The evaluation focuses on five key aspects of feedback quality: adherence to criteria, clarity of suggestions for improvement, accuracy, prioritization of essential writing elements, and delivery in a supportive tone. The comparison of descriptive statistics and effect sizes revealed that well-trained human raters demonstrated higher quality feedback compared to ChatGPT. Despite this finding, the accessibility and generally acceptable quality of feedback obtained through ChatGPT suggest that generative AI could remain beneficial in specific contexts, such as for initial drafts or in situations where access to highly trained raters is limited.

Online Profile

Biography

Primary author: DO TRONG, Hoang (University of Languages and International Studies)

Presenter: DO TRONG, Hoang (University of Languages and International Studies)

Session Classification: Posters

Track Classification: Teaching Methods and Applied Linguistics