British Journal of Applied Linguistics
ISSN: 2754-5253
DOI: 10.32996/bjal
Journal Homepage: www.al-kindipublisher.com/index.php/bjal

BJAL
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Human vs ChatGPT Feedback in EFL Writing: A Comparative Study of Error Correction and Learner Uptake Among Saudi University Students

Ali S. Alghonaim
Associate Professor, Department of English Language and Literature, College of Languages and Humanities, Qassim University, Buridah, Saudi Arabia
Corresponding Author: Ali S. Alghonaim, **E-mail**: agniem@qu.edu.sa

| ABSTRACT
The rapid spread of artificial intelligence (AI) in language education has led teachers to use this tool in teaching and pedagogical practices across several disciplines. In English as a foreign language (EFL) context, various AI tools are being applied in teaching practices. ChatGPT is one of the most popular tools currently used by both teachers and students. This research compares human feedback and ChatGPT feedback as two methods used in Saudi EFL settings. The study adopted a quasi-experimental design. The learner participants were divided into two groups: a human feedback group and a ChatGPT feedback group. The research focused on different writing concerns addressed by the two methods of feedback. These feedback notes include grammar, vocabulary, organization, and discourse. Also, the study focused on the different responses of the students to the two types of feedback. The results revealed that ChatGPT feedback provided instant feedback on the surface level of the language structures. Human feedback, however, treated higher levels of contextual writing concerns. The study also focused on the students' rates of response to the two types of feedback. The results showed that the ChatGPT feedback group revised their grammatical mistakes rapidly. However, the students in the human feedback group spent a longer time dealing with contextual issues. Overall, the main recommendation of this research is that both types of feedback are helpful in improving EFL writing.

| KEYWORDS

Artificial intelligence, ChatGPT, EFL writing, written corrective feedback, human feedback, learner uptake, Saudi Arabia

| ARTICLE INFORMATION

ACCEPTED: 01 April 2026 **PUBLISHED:** 19 April 2026 **DOI:** 10.32996/bjal.2025.6.2.2

1. Introduction

Writing is often considered one of the most challenging skills for EFL students because students need to control a number of areas at once, such as grammar, vocabulary, organization, and overall coherence (Hyland, 2003). Therefore, feedback is very important because students may notice their mistakes and improve their writing over time (Ellis, 2009).

It has been commonly noticed that most EFL classrooms teachers represent the main source of feedback (Ferris, 2014). In fact, limited class time and large numbers of classes represent serious difficulties in dealing with students' writing. Additionally, subjectivity plays another challenge as it allows for differences in teachers' reaction and feedback to their students writing assignments (Bitchener & Knoch, 2008).

Recently, writing teachers use tools such ChatGPT in order to correct and provide feedback. According to Godwin-Jones, 2022; Mao et al., 2023, students receive instant and immediate response from their teachers. However, such tool is still under

investigating whether it can fully deal with deeper writing issues, such as meaning, organization, and whether it downgrade itself to the students' writing levels (Ellis, 2012).

In Saudi Arabia, English writing courses are enrolled in many English department plans in most Saudi universities. This shows that this skill represents an important academic skill. Therefore exploring if feedback provided by using ChatGPT is necessary to understand whether this ChatGPT-generated feedback is really as useful as traditional teacher feedback.

2. Literature Review

2.1 Feedback in Second Language Writing

Feedback is widely recognized as essential in second language acquisition and writing development. It helps learners notice gaps between their output and target language norms (Schmidt, 1990; Ferris, 2011). Corrective feedback may be direct or indirect and plays a key role in interlanguage development (Ellis, 2006).

Although technology advancement provided writing teachers with different tools and aids to be employed in writing correction, teachers still maintain the normal practice of correction by giving written corrective feedback. Teachers believe that teacher corrective feedback recognizes the contextual factors surrounding learners' writing and understand pedagogical flexibility (Hyland & Hyland, 2006). The main obstacle, however, is the time teachers spend to correct, especially with large classes (Chaudron, 1988).

2.2 Learner Uptake in Writing Feedback

Lyster and Ranta, 1997 identify learner uptake as learners' reaction to the corrective feedback provided by their writing teachers and to what extent they respond to the teacher's notes and comments. Nassaji, 2015 argues that corrective feedback may require writers to change some parts of their work rewrite the whole task and that learners' reaction depends on some factors such as, learners level of writing proficiency, aptitude, nature of the writing task and clarity of feedback. As mentioned above, writing teachers prefer to use written feedback. Therefore, most research focuses on teacher feedback and give limited attention to ChatGPT-generated feedback.

2.3 AI in Language Learning and Writing Feedback

Although teachers still maintain the use of human corrective feedback, they start adopting Artificial intelligence (AI) in language teaching, in general, and in corrective feedback, in particular (Zawacki-Richter et al., 2019). ChatGPT is the easiest means tool as it is considered a popular platform and provides immediate responses. It also generates other options of feedback corrections (Godwin-Jones, 2022).

However, although this tool helps develop grammatical accuracy and supports learner autonomy, previous research on AI, in general, and on ChatGPT, in particular, shows that it may not provide deeper pedagogical explanation and does not accommodate the discourse-level correction (Mao et al., 2023). This raises a question that ChatGPT primarily focuses on surface linguistic forms rather than on meaning construction (Ellis, 2012).

2.4 Human vs AI Feedback in EFL Writing

Human feedback deals with human writing and therefore it is gives a special attention to discourse-level and naturally addresses other aspects of writing other than the grammatical forms. It assesses meaning creation, coherence, and structural organization and rhetorical functions (Ferris, 2004; Hyland, 2019). However, AI feedback renders instant responses consistent, objective, and computable (Godwin-Jones, 2022).

Despite growing interest in utilizing AI in language learning and teachings, few empirical studies directly study human feedback and AI feedback in EFL writings.

2.5 Research Gap

The review of literature above proposes two main research gaps that show the significance of this research. The first gap relates to the limited or few studies that focuses on the comparison between human feedback and ChatGPT-generated feedback in EFL academic writing, in general, and Saudi EFL writing learners. In particular. The second research gap, according to Lyster and Ranta, 1997, is how learners respond to ChatGPT-automated feedback compared to human feedback.

3. Research Questions

1. Do human and ChatGPT-generated feedback respond to errors in similar effectiveness?
2. How do learners uptake differ between ChatGPT and human feedback?
3. Which feedback type leads to greater improvement in EFL writing?

4. Methodology

This research used A quasi-experimental design. Pretest-posttest writing tasks were used to compare between human teacher feedback and artificial intelligence platfrom (ChatGPT). Specifically, the study aims to explore which strategy statistically enhances EFL writing more.

4.1 Participants, Instruments and Procedures

The participants were 36 undergraduate Saudi EFL students studying in the English Department at Qassim University. Their age rages between 18-25 years old. 18 randomly selected students received human feedback (control group and 18 students received ChatGPT generated feedback (experimental group). The two groups completed a Pre-test and a psot-test writing task of 200 -300 word paragraphs. This method allows the researcher to compare between the two groups. The research used a rubric tool to assess the learners' writing. This rubric tool is developed by Ferris, (2014). The writing was analyzed for grammatical, lexical, structural, and discourse errors. The analytical scoring was as follows:
-Grammar (10)
-Vocabulary (10)
-Organization (10)
-Coherence (10)
Total score: 40 points

The study used descriptive statistics (mean and standard deviation), paired samples t-test, and Cohens' d to measure the effect size, the significant level was set at a =0.05

5. Results

5.1 Descriptive Statistics

Group	Test	Mean	SD
Human	Pre	21.3	3.2
Human	Post	30.8	3.6
AI	Pre	21.7	3.1
AI	Post	27.2	3.4

Both groups improved from pre-test to post-test, with greater gains observed in the human feedback group.

5.2 Within-Group Improvement

Human group: t(17) = 9.21, p < 0.001, d = 1.85.
ChatGPT group: t(17) = 6.14, p < 0.001, d = 1.10.
Both groups showed statistically significant improvement.

5.3 Between-Group Comparison

t(34) = 2.87, p = 0.007, d = 0.95.
The human feedback group significantly outperformed the AI group.

5.4 Error Analysis

Error Type	Human (%)	ChatGPT (%)
Grammar	78	90
Vocabulary	85	70
Organization	88	65
Discourse	90	60

5.5 Learner Uptake

Response Type	Human (%)	ChatGPT (%)
Full	52	72
Partial	38	22
None	10	6

ChatGPT feedback led to higher full uptake, while human feedback encouraged more partial uptake.

5.6 Writing Performance

This current research adopted descriptive statistical analysis of the results of the pre-tests and post-tests to compare between teacher feedback and CharGPT- generated feedback.

The result of the two groups in the pre-test and post-test indicated noticeable improvement in the human teacher feedback and the ChatGPT-generated feedback. learners scored M = 21.3 (SD = 3.2) in the pre-test of the hhuman feedback compared to a mean score of M = 30.8 (SD = 3.6) in the post-test. The score of the ChatGPt-generated feedback showed similar improvement. The learners scored M = 21.7 (SD = 3.1) in the pre-test and M = 27.2 (SD = 3.4) in the post-test.

A paired-samples t-test revealed that the improvement in the human feedback group was statistically significant, **t(17) = 9.21, p < .001**, with a **very large effect size (d = 1.85)**. Likewise, the ChatGPT feedback group also showed a statistically significant improvement, **t(17) = 6.14, p < .001**, with a **large effect size (d = 1.10)**.

As shown above, the two groups achieved progress from the two kinds of feedback. However, the human feedback group achieved relatively higher improvement. Independent-sample t-test of the post-test scores confirmed the outperformance of the human feedback group over the ChatGPT feedback group, t(34) = 2.87, p = .007, with a large effect size (d = 0.95).

5.7 Analysis of Writing Components

ChatGPT-generated feedback benefited from feedback in the surface level. For example, the findings revealed that students in this group improved in grammar with an average of score of +8.2, compared to +6.5 points in the human feedback group. Clearly, this feedback group gained effective notes and comments about rule-based language structures.

However, the two groups differed in higher-level writing elements. The human feedback showed better improvement in vocabulary (+5.1 vs. +3.4), organization (+4.8 vs. +2.1), and discourse and coherence (+5.2 vs. +1.9). Yet, the findings showed that ChatGPT was provided sufficient feedback on the surface level. The deeper structural features of writing seemed to be effectively addressed by human feedback because such writing aspects require contextual awareness that ChatGPT is still far away at this moment.

5.8 Error Correction Patterns

The result revealed statistical data about the correction rates that support the findings of the study. Correction rate in grammar was higher in ChatGPT feedback. Students in this group achieved 90% compared to 78% achieved by the human teacher feedback group. However, human feedback group achieved higher in:

- **Vocabulary (85% vs. 70%)**
- **Organization (88% vs. 65%)**
- **Discourse (90% vs. 60%)**

The differences were statistically significant ($p < .05$). This indicates that feedback uptake depends on the type of errors.

5.9 Learner Uptake

The assumption that learners may vary in their response to feedback was among the focus of this research. Therefore, learner uptake was investigated to examine learners' reaction to the provided feedback. The findings showed that the two groups responded differently. Students in the human feedback group achieved a lower rate of full uptake (52%), compared to a higher rate of full uptake (72%) in the ChatGPT feedback group. However, the findings showed that the ChatGPT achieved lower rate of partial uptake (22%) compared to (38%) in the human feedback group. Finally, the two groups relatively presented low rates of no uptake with slight difference in favor of the human feedback (10%) and (6%) for the ChatGPT feedback.

The findings regarding uptake patterns were significant ($p<.05$). to interpret these findings, human feedback promotes reflective reaction and response to the writing task, whereas, ChatGPT feedback leads to quick revision and immediate changes.

Discussion and conclusion

The findings of this study support results of previous research on feedback in EFL writing. For example, the effectiveness of human feedback on higher-level writing features in this study aligns with research study conducted by Ferris (2014). The two studies confirm that some writing features, such as coherence, rhetorical organization, discourse develop more through human-related context awareness.

Conversely, ChatGPT feedback provided more effective grammatical notes. This is supported by Jones, (2022) and Mao et al. (2023) who report that Chat-GPT-generated feedback outperform traditional teacher feedback in on the surface levels of language structures. Ellis (2012) argues that technology still lags behind in the interpretation of the contextual meanings of words and expressions.

Rates of learner uptake in the study indicate differences between the two feedback groups. The framework suggested by Lyster and Ranta (1997) accounts for the variation of the leaner uptake; the ChatGPT feedback group achieved higher rate of full uptake because they received the feedback immediately after they finish the writing task. On the other hand, the human feedback group achieved higher rate of partial uptake. This may be traced back to higher level of engagement with feedback. This supports what Nassaji's (2015) found earlier. He pointed out that more processing leads to successful long-term learning.

So far, this research finds evidence that supports previous studies on feedback in EFL writing. However, it also adds some contributions to the literature. Specifically, it addresses the gaps identified in the literature review section. It compares human and ChatGPT-generated feedback in a Saudi EFL context. A further contribution of this research is the investigation of the different levels of learners' uptake of feedback from the two methods.

Pedagogically speaking, a compromise can be adopted to deal with human feedback and ChatGPT feedback. The research results indicate that the two methods of feedback complement one another. The ChatGPT feedback method works better for instant grammatical accuracy, whereas human feedback addresses deeper writing concerns such as genre, structural organization, coherence, and discourse.

Using ChatGPT, or any Artificial Intelligence (AI) feedback tool, may be a useful technique for providing feedback to students in large class sizes, as is the case in most Saudi higher education settings. These AI tools should assist teachers in delivering feedback while maintaining teaching quality. The findings in favor of human feedback indicate that absolute reliance on AI may not be sufficient to address advanced aspects of writing.

To conclude this research, a balanced approach between AI and human feedback is needed. Although AI feedback is useful, human intervention is indispensable for dealing with human linguistic behaviors. Human feedback appears to be highly sensitive to discourse and contextualized writing.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1]. Alexopoulou, T., et al. (2023). AI-assisted writing feedback in language learning: Opportunities and challenges. *ReCALL, 35*(3), 1–18.
[2]. Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing, 17*(2), 102–118.
[3]. Bitchener, J., & Ferris, D. (2012). *Written corrective feedback in second language acquisition*. Routledge.
[4]. Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback. *Language Teaching Research, 12*(3), 409–431.
[5]. Brown, H. D. (2007). *Principles of language learning and teaching* (5th ed.). Pearson.
[6]. Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. Cambridge University Press.
[7]. Ellis, R. (2006). Researching the effects of form-focused instruction. *Studies in Second Language Acquisition, 28*(2), 139–186.
[8]. Ellis, R. (2009). *Corrective feedback and teacher development*. L2 Journal, 1(1), 3–18.
[9]. Ellis, R. (2012). *Language teaching research and language pedagogy*. Wiley-Blackwell.
[10]. Ferris, D. (2004). The "grammar correction" debate revisited. *Journal of Second Language Writing, 13*(2), 49–62.
[11]. Ferris, D. (2011). Treatment of error in second language student writing. *Annual Review of Applied Linguistics, 31*, 45–63.
[12]. Ferris, D. (2014). *Written corrective feedback in second language acquisition and writing*. Routledge.
[13]. Godwin-Jones, R. (2022). Emerging technologies: ChatGPT and language learning. *Language Learning & Technology, 26*(2), 1–5.
[14]. Hyland, K., & Hyland, F. (2006). *Feedback on second language students' writing*. Cambridge University Press.
[15]. Hyland, K. (2003). *Second language writing*. Cambridge University Press.
[16]. Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press.
[17]. Krashen, S. (1985). *The input hypothesis*. Longman.
[18]. Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of L2 development*. Oxford University Press.
[19]. Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition*.
[20]. Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition, 19*(1), 37–66.
[21]. Mackey, A. (2007). *Input, interaction and corrective feedback in L2 learning*. Oxford University Press.
[22]. Mao, Z., et al. (2023). Artificial intelligence in EFL writing: A systematic review. *Computers and Education: Artificial Intelligence, 4*, 100–120.
[23]. Nassaji, H. (2015). Interactional feedback and SLA. *Language Teaching Research, 19*(2), 137–160.
[24]. O'Keeffe, M., McCarthy, M., & Carter, R. (2007). *The Routledge handbook of corpus linguistics*. Routledge.
[25]. Polio, C. (2012). The relevance of second language writing research. *Language Teaching, 45*(3), 261–279.
[26]. Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158.

[27]. Swain, M. (1985). Communicative competence: Some roles of comprehensible output. In S. Gass & C. Madden (Eds.), *Input in second language acquisition*.
[28]. Swain, M. (2005). The output hypothesis. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning*.
[29]. VanPatten, B. (2004). *Processing instruction*. Routledge.
[30]. Vygotsky, L. S. (1978). *Mind in society*. Harvard University Press.
[31]. Zawacki-Richter, O., et al. (2019). Systematic review of AI in higher education. *International Journal of Educational Technology in Higher Education, 16*(1), 39.

Appendixes

Appendix A- TABLES

✓ **Table 1: Participant Design**

Group	Type of Feedback	N	Description
A	Human teacher feedback	18	Written corrective feedback
B	feedback (ChatGPT)	18	Automated corrective feedback

✓ **Table 2: Error Classification Framework**

Error Type	Description	Example
Grammar	tense, agreement	*He go → He goes*
Vocabulary	word choice	*make a photo → take a photo*
Organization	structure	paragraph order issues
Discourse	cohesion/coherence	weak linking

✓ **Table 3: Results Summary**

Measure	ChatGPT Feedback	Human Feedback
Grammar correction	High	Medium
Vocabulary improvement	Moderate	High
Organization improvement	Low	High
Learner uptake	Immediate	Reflective

✓ **Table 4: Learner Uptake**

Uptake Type	ChatGPT Group	Human Group
Full uptake	High	Moderate
Partial uptake	Moderate	High
No uptake	Low	Low

✓ **Appendix B — Writing Task Prompt**

Example:

Write a 150–200 word paragraph about "The importance of technology in education."
Focus on grammar, vocabulary, and organization.

✓ **Appendix C — ChatGPT Feedback Sample**

Example:

Student sentence:

He go to university every day.

ChatGPT feedback:

Correct form: *He goes to university every day.*
Explanation: Use third-person singular present simple rule.

✓ **Appendix D — Teacher Feedback Sample**

Student sentence:

He go to the university every day.

Teacher feedback:

Correct verb form (subject-verb agreement)

Compare between "to the university" and " to university"