ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

A MIXED-METHODS ANALYSIS OF TEACHERS' AND STUDENTS' PERCEPTIONS OF NLP-GENERATED WRITING FEEDBACK USING COH-METRIX

Arslan Anjum
MS Scholar, Applied Linguistics, National University of Computer and Emerging Sciences, Lahore, Pakistan.
arslananjum1098@gmail.com
Arzeen Bhatti
MS Scholar, Applied linguistics, University of Management and Technology, Lahore, Pakistan.
arzeenbhatti4@gmail.com
Ayesha Imtiaz
Adjunct Faculty FAST-NUCES, MPhil Applied Linguistics from Kinnaird College for Women, Lahore, Pakistan.
imtiazayesha031@gmail.com

Abstract

This study investigates teachers' and students' perceptions of NLP-generated writing feedback produced through Coh-Metrix, a computational linguistic tool developed by Graesser and McNamara (2004). Although NLP feedback systems are increasingly integrated into educational contexts, little is known about how learners and instructors interpret, trust, and utilize such feedback in real writing situations. To address this gap, the study employs a mixed-methods design combining computational text analysis with qualitative interview data. Writing samples produced by students were analyzed using Coh-Metrix to generate indices of cohesion, lexical sophistication, syntactic complexity, and readability. Semi-structured interviews were conducted with both students and teachers, and the resulting transcripts were examined through Braun and Clarke's (2006) thematic analysis. Quantitative outputs from Coh-Metrix were compared with participants' perceptions to identify areas of alignment and mismatch between automated evaluations and human judgments. Findings are expected to reveal how the participants interpret NLP feedback, the extent to which they trust computational assessments, and the challenges they face when integrating such feedback into writing instruction or revision practices. The study contributes to a deeper understanding of how NLP tools can be effectively implemented in educational settings and provides insights for enhancing the pedagogical usefulness of automated writing feedback systems.

Keywords: Natural Language Processing; Coh-Metrix; Computational Text Analysis; Automated Evaluation

Introduction

NLP technologies have revolutionized the manner in which writing is taught, assessed and assisted. Writing feedback tools have become prominent especially automated ones, which give instant linguistic feedback about the written texts of learners. One of these tools is Coh-Metrix, which has the ability to measure cohesion, lexical sophistication, and syntactic complexity using the computational mechanisms (Graesser et al., 2004).

Since more educational environments are embracing AI-driven feedback systems, the perception of these systems as either beneficial or harmful to the educational process emerges as an important field of investigation. Application of NLP tools in the writing practices has demonstrated the possibility of possible improvement of the learner autonomy and writing quality but the acceptance of the tools in pedagogy is not uniform. Automated writing evaluation (AWE) has been discovered to provide valuable feedback, yet students tend to doubt the direction and precision of the feedback, especially when it opposes teacher feedback (Link et al., 2020).

Educators can also also be concerned with the pedagogical suitability of machine-generated remarks or the danger of simplified explanations of the complexity of language. Such opposing perspectives indicate the necessity of the empirical research that will compare the computational feedback to the human perceptions. Simultaneously, there is considerable literature that points to the significance of feedback as a form of social and cognitive processing predetermined by the interpretation and belief of the writers (Hyland and Hyland, 2006). When the source of feedback is an NLP system, not a human teacher, the learners and instructors will react differently to these sources because of the trust, familiarity, or perceived usefulness.

Technology Acceptance Model (TAM) is based on the idea that ease of use and perceived usefulness have a mighty effect on the acceptance of educational technologies by users (Davis, 1989). Thus, the implementation of the TAM to the NLP-generated writing feedback situation can help understand whether these tools can meet the expectations and pedagogical objectives of the users. Recent improvements in writing support systems based on AI made them more attractive to educational applications due to the accuracy of the linguistic features detection. Nevertheless, research shows that automated feedback might be difficult to make sense without the help of instruction (Zhang, 2023).

Due to the fact that computational feedback is aimed at text features and not communicative intent, teachers need to mediate the process of including such feedback in the process of teaching classroom writing. The dynamics of these tools would only be understood after a qualitative investigation of the real-life experiences of users using NLP tools. Although NLP feedback tools are used with more frequency, no mixed-method studies to evaluate the comparative effectiveness of computational evaluations versus the lived experiences and perceptions of both teachers and students exist.

Although Coh-Metrix provides advanced results of textual quality, its worth is subject to the perceptions and applications of its feedback by the user during the writing and revision process. This research study aims at filling the gap between the computational representations of text quality and human understandings of the writing effectiveness through the integration of Coh-Metrix analysis and interview data. These observations can help enhance the adoption of AI-based feedback instruments in learning environments.

Research Questions

1. What is the perception of teacher and student concerning the usefulness, accuracy and relevance of the writing feedback produced by Coh-Metrix?
2. In what ways do the computational features, found in Coh-Metrix, compare or contrast with the qualitative interpretation of writing quality done by teachers and students?

Objectives

1. To analyze teachers and student perceptions of NLP generated writing feedback created with Coch-Metrix.
2. To draw comparison between Coh-Metrix computational text analysis findings and human ratings of quality writing to determine points of intersection and divergence.

Literature Review

With the growing adoption of Natural Language Processing (NLP) as a part of the learning process, the way writing is learned, judged, and facilitated has changed. The automated scoring systems and AI-based linguistic analysis systems represent some of the opportunities provided by the technologies in writing development. To be effective, however, the interpretations and engagements of computational feedback as perceived by learners and teachers are vital. Recent research indicates the potential of NLP-based writing aids as well as the issues of the pedagogical application of these applications, and it has shown that further investigation of the perception-based study is necessary to compare the human judgement with the automated output.

1. NLP Tools and Their Role in Writing Development

In the past few years NLP tools have developed to a great extent, providing in-depth linguistic analytics that can assist learners in gaining a clearer understanding of their writing performance. Automated feedback systems and linguistic analyzers are tools that may assist in instant feedback regarding grammar, cohesion, and lexical peculiarities and enable more effective revision of the work (Lu et al., 2021). Such tools are also beneficial in enabling the students to detect their lingual issues that would otherwise be invisible.

Such benefits notwithstanding, studies have established that NLP feedback can only be effective in cases where the learners know how to decode them. Automated suggestions might be misunderstood or learners might fail to make substantial revisions without instructions (Stevenson et al., 2023). This implies that the use of NLP tools requires the inclusion of instructional support to achieve the full effects of the tools on the quality of writing.

2. Automated Writing Evaluation Systems

AWE systems have become popular since they offer scoring and feedback as well as linguistic analysis that once took a lot of teacher time to obtain. The current AWE systems use machine-learning algorithms to estimate the coherence, lexical sophistication, and syntactic complexity (Zhang and Hyland, 2022). Their speed and accuracy have rendered them to be more practical in the classroom.

Yet, not even state-of-the-art AWE systems can be limited to nothing. It has been noted that teachers feel that automated tools cannot assess the discourse-level features of the quality of argumentation and the rhetorical structure (Ranalli and Arrigoni, 2019). It points to the persistence of the necessity to study the correspondence and lack of correspondence between computational assessments and human expertise.

3. Coh-Metrix and Computational Text Analysis

Coh-Metrix is also very popular in terms of writing research because it has the capability of creating indices concerning cohesion, syntactic complexity, and lexical difficulty. It provides more profound linguistic knowledge than the error-detection task on the surface, which is why it is helpful when it comes to assessing the progress in the field of writing (McNamara et al., 2017).

The multidimensionality of the tool allows scholars to have extensive linguistic profiles of student texts. However, there is a difference between computational linguistic measures and what users think of the quality of writing. It has been established that Coh-Metrix scores are not necessarily consistent with teacher ratings or student views of their writing strengths (Crossley et al., 2018). This gap necessitates a need to find ways of integrating human and automated assessments.

4. Teacher Perceptions of Automated Feedback

The perceptions of the teachers are also a critical factor in deciding the manner in which NLP tools are employed in teaching writing. Other teachers consider automated feedback to be useful to complement classroom learning and provide students with additional revision opportunities (Li et al., 2019).

Those teachers, who find NLP tools to be correct and pedagogically orientated, are more inclined to implement them successfully (Khan et al., 2025). But some people also raise their concerns about the reliability and pedagogical importance of machine-generated suggestions. The doubts about automated feedback shared by many educators are that it could simplify some of the complex writing constructs or demean the role of the teacher in formative assessment (Wilson and Roscoe, 2020). These conflicting views indicate the significance of investigating the teacher attitudes towards the NLP-generated feedback.

ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

5. Student Engagement with Automated Feedback

The interaction of students with automated feedback determines NLP effectiveness significantly. Studies show that automated feedback is typically valued by learners as it contributes to the increased frequency of the revision and self-directed learning (Bai and Hu, 2021). In a case whereby students have confidence in automated recommendations, then they are bound to take significant action in enhancing their writing.

Conversely, there are students who doubt the correctness or usefulness of automated feedback, particularly when it does not match the evaluations of teachers or when they provide vague recommendations (Lee et al., 2021). These impressions influence the way the students will incorporate the NLP feedback in their process of writing; thus, a greater understanding of the learner perceptions is a requirement.

Research Methodology

The research design used in the current study is a mixed-methods one as it incorporates both qualitative or quantitative elements to offer in-depth insights into the nature of the perceptions of both the teachers and students towards NLP-generated writing feedback. A mixed-methods research is particularly appropriate since it focuses on (a) computational text analysis in Coh-Metrix and (b) the perceptions of participants to be obtained in interviews. In alignment with the explanatory sequential model described by Creswell and Plano Clark (2018), the analysis of writing samples is conducted by the researcher in a quantitative manner with the help of the Coh-Metrix indices and then, by a qualitative method organized as a thematic analysis that delineates the causal relationships in explaining and supplementing the computational results. Such a progressive integration enables strong triangulation and more insights than a unilateral approach.

1. Participants and Sampling

The respondents will consist of 10-15 English writing classes students and 5-10 English language teachers who have taught writing. Participants are recruited by purposive sampling, criteria being that the sample consists of people who were previously exposed to writing tasks or are conversant with automated feedback tools. This sampling will be suitable since participants will be able to give deep, informative information that is in line with the purpose of the study. The sample size of mixed-methods inquiry is also decent to achieve saturation in qualitative data and sufficient representation to interpret data quantitatively.

2. Research Instrument

2.1. NLP Tool: Coh-Metrix

The major quantitative tool is Coh-Metrix created by Graesser and McNamara (2004). It produces precise linguistic indices on the topics of cohesion, syntactic complexity, lexical sophistication, and readability. Such computational results aid objective assessment of writing characteristics of students.

2.2. Semi-Structured Interview Protocol

Both teachers and students are interviewed using semi-structured interviews in order to determine their views on Coh-Metrix-generated feedback. The interview questions are created in such a way that they will help to obtain the perception of the respondents concerning accuracy, clarity, usefulness, and correspondence to their expectations of writing feedback.

3. Data Collection Procedure

Step 1: Recruitment and Consent

The first step is the recruitment and consent. Institutional recruiting is used to recruit the participants and the informed consent is written. The intent, secrecy and voluntary-ness of participation are clearly explained.

Step 2: Writing Sample Collection

Students write a short writing activity (e.g., an essay or a paragraph), which becomes an input to Coh-Metrix analysis.

Step 3: Coh-Metrix Analysis

Writing samples of the students are sent to Coh-Metrix in order to produce quantitative linguistic indices. These are the outcomes of the calculation assessment dataset.

Step 4: Interviews

Individual semi-structured interviews with teachers and students are made. Interviews are taped to enable accuracy in the course of transcription.

Step 5: Transcription and Verification

Every record of the interviews is transcribed word-to-word. Participants are encouraged to confirm the truthfulness of their answers, which makes them credible by member checking.

4. Data Analysis

Mean and standard deviations are used as descriptive statistics to analyze the Coh-Metrix output. The metrics can be used to determine the trends of cohesion, lexical complexity, syntactic complexity and readability in writing samples. Interview transcripts are subjected to Braun and Clarke (2006) six phases of thematic analysis familiarization, coding, theme generation, theme refinement, definition, and reporting. This secures systematic and rigorous sensemaking of the perceptions of the participants. In accordance with the mixed-methods integration strategy suggested by Creswell and Plano Clark (2018), the results of Coh-Metrix are contrasted with the patterns that appear after the interview. The interpretation of convergences and divergences is made to create one coherent view of the way of how computational feedback and human judgments meet.

5. Theoretical and Analytical Framework

This study is guided by three key theoretical and analytical frameworks:

5.1. Technology Acceptance Model (TAM) -Davis (1989)

The Technology Acceptance Model assumes that the acceptance of technological tools by users is determined by the perceived usefulness and perceived ease of use. TAM offers a theoretical prism through which students and teachers can be viewed in terms of acceptance or resistance to NLP-generated feedback in terms of perceived accuracy, clarity, and practicality.

5.2. Feedback Theory Hyland and Hyland (2006)

Feedback has already been conceptualized as a social and cognitive process with a focus on the way learners perceive and react to evaluative comments. This framework assists in the interpretation of the way participants negotiate meaning when machine-generated feedback is not in line with the expectations of the teachers.

5.3. Computational Linguistic Framework -Graesser et al. (2004)

This model aids the analytical aspect of the research provided to explain the way Coh-Metrix measures linguistic characteristics like cohesion and syntactic complexity. It enables the making of computational indices to be logically interpreted and the qualitative perceptions. The combination of these frameworks allows not only the interpretation of the perceptions of participants but the analysis of computational linguistic features which guarantees integrity in the methodological sense.

6. Ethical Considerations

Ethical consent is received before data is collected. The objectives of the study, the possible risks and confidentiality of the study as well as the participants right to opt out of the study at any point is explained to them. Participants have identities that are safeguarded by pseudonyms. Tapes are well secured and can only be accessed by the research team.

ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

Results and Findings

This part is a summary of the most important findings of the Coh-Metrix analysis of the writing samples of students, as well as the most dominant themes of the interviews with teachers and students. The results are grouped into quantitative and qualitative aspects so as to indicate the compatibility of the computational linguistic attributes with the human perception of writing feedback. The subsections consist of tables that summarize the data and then present the interpretations that explain the relevance of the results as per the research questions of the study. Scores obtained through Cohesion Co-Metrix.

1. Cohesion Scores Generated by Coh-Metrix

Table1

Cohesion Metrics for Student Writing Samples

Student ID	Referential Cohesion	Deep Cohesion	Latent Semantic Similarity
S1	0.42	0.31	0.55
S2	0.51	0.44	0.61
S3	0.38	0.29	0.48
S4	0.47	0.36	0.58
S5	0.53	0.41	0.62

Coh-Metrix coherence findings depict moderate rates of referential and deep cohesion among writing samples of students. The best scores in cohesion were obtained by the students S2 and S5, which means that they used connective devices and logical flow to a greater extent. Reduced scores in deep cohesion (e.g., S3) provide indications of difficulty in incorporating cause or purposeful connections in text. On the whole, the results show that student ability to relate ideas varies, which emphasizes the need to take a more in-depth look at the areas of potential positive outcomes of explicit instruction on cohesion.

2. Syntactic Complexity Measures

Table2

Syntactic Complexity Scores

Student ID	Mean Sentence Length	Left-Embeddedness	Complex Nominals per Clause
S1	12.4 words	0.18	0.35
S2	15.2 words	0.22	0.41
S3	10.8 words	0.16	0.30
S4	14.6 words	0.20	0.39
S5	16.1 words	0.23	0.44

The outcomes of the syntactic complexity point to the fact that students tend to have moderate sentence sophistication. S2 and S5 students are more complex, and they write longer sentences with more complex structures built in them. Reduced complexity score by students such as S3 implies that they rely on the simple forms of sentences. The difference shows that though there are learners who make more attempts to construct more complicated grammar structures, there are learners who need more help in syntactic development.

ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

3. Students' Perceptions of NLP-Generated Feedback

Table3

Summary of Student Interview Themes

Theme Identified	Number of Students Mentioning It	Example Expressions
Usefulness of Immediate Feedback	11	"It helps me revise faster."
Confusion About Technical Terms	8	"I don't understand some linguistic words."
Trust in Accuracy	6	"Sometimes it feels wrong or unclear."
Motivation to Revise	9	"It makes me revise more willingly."

Students were favorably impacted by the instantaneity and convenience of Coh-Metrix feedback with 11 students mentioning it as helpful in revisions. Nonetheless, misunderstanding of technical lingo language terms was a common issue, which means that teacher mediation was necessary. Most students liked the guidance of the tool but only half of them indicated that they highly trusted it to be in the right way. The tool itself seemed to have a stimulative effect on revision behavior, with motivational returns even though it was not easy to interpret it.

4. Teachers' Perceptions of NLP Feedback

Table4

Summary of Teacher Interview Themes

Theme Identified	Number of Teachers Mentioning It	Example Expressions
Helpful for Initial Diagnosis	7	"Good for first-round checking."
Limited in Evaluating Argumentation	6	"It cannot judge ideas or logic well."
Concern Over Student Overreliance	5	"Students may depend on it too much."
Need for Teacher Mediation	8	"It requires explanation and context."

The overall opinion of the teachers about Coh-Metrix was positive as it helped them to conduct a primary assessment of the linguistic background especially in cases of any deficiency in cohesion and complexity. Nonetheless, six educators pointed out the weakness of the tool to assess higher-order writing components like quality and coherence of the argument. A lot also raised the concern that automated feedback will make students too reliant on it. The most repeated theme is teacher mediation, which shows that teachers consider NLP tools as helpful, but not as substitutes to human feedback.

Discussion

This discussion aims at interpreting the research findings on the basis of existing research and research questions that were used to guide the study. The discussion, by looking at the Coh-Metrix analysis and perceptions of participants, shows how the computational feedback agrees or disagrees with those of human evaluations of writing. It also discusses the pedagogical consequences of applying NLP tools in writing instruction with a specific focus on the strengths and limitations that the data produced.

1. Interpretation of Cohesion Findings in Relation to Student Writing Development

Coh-Metrix scores on cohesion reveals that even though some students exhibit moderate scores in referential and deep cohesion, others have difficulties in sustaining logical links in their texts. Higher-scoring students including S2 and S5 seem to write with more semantic connections that are consistent and this seems to indicate that they are more familiar with cohesive devices. Such inconsistency in the cohesion is consistent with previous studies that have found that learners tend to be very different with regard to their capacity to build coherent discourse (Crossley et al., 2019). The reduced deep cohesion scores of such students as S3 could be interpreted as the difficulties of assimilation of causal and intentional relations, which justifies the necessity to teach the discourse-level coherence explicitly. The findings bring to light that much as NLP tools such as Coh-Metrix may be able to find out patterns of cohesion, human interpretation is essential in putting them into perspective in terms of the writing growth of individual learners.

2. Syntactic Complexity and Students' Linguistic Awareness

The results of syntactic complexity show the obvious dissimilarity of the sentence structure and grammatical complexity of students. Students with overall writing complexity were also found to have more structure embedded in their sentences and thus lengthy sentences, which is in line with the findings of Lu (2021) that syntactic variety is related to more advanced writing. On the other hand, students who have shorter and simpler sentences might not be confident or aware of the ways of using complex grammatical forms. These findings depict that Coh-Metrix is efficient in observing linguistic variations, yet it fails to state why students prefer this or that structure. The results of interviews with students have shown that some of them did not know how their syntactic decisions affected writing quality, there was a disconnection between computational feedback and metalinguistic knowledge of students. Therefore, mediation by the teacher is necessary in assisting learners to read and use complexity scores and use them in revision.

3. Students' Perceptions of NLP-Generated Feedback

All in all, the students were positive about the instantaneous and the availability of automated feedback. This is consistent with the research that automated systems bring more motivation to revising among learners (Bai and Hu, 2021). Nevertheless, ambiguity in language terminology implies that learners can be helped through the scaffolded teaching on computational feedback. Although a large percentage of students also found the feedback valuable, only a fraction of them expressed that they trusted the accuracy of the feedback very much, which, again, is not novel, as past studies show that students tend to doubts the accuracy of the automated assessments (Li et al., 2019). The inspirational feature of NLP feedback is positive yet it is evident that unless the students are coached, they can interpret the information given or abuse it. It implies that NLP feedbacks cannot act independently and therefore they must be combined with teacher explanations in order to have the best learning outcomes.

4. Teachers' Perceptions and the Need for Instructional Mediation

The teachers were found to mostly love Coh-Metrix as a diagnostic instrument that assists in detecting the preliminary linguistic problems in student writing. Nevertheless, their worries about the constraints of the tool, particularly in terms of argumentation and development of ideas, are the same as those of Wilson and Roscoe (2020), who observe that automated systems do not cope with high-level writing. The concern expressed by teachers over the overdependence of the students is indicative of a larger debate of technological support and student independence. Although teachers did not disapprove of the role of the tool in analysis of languages, they always insisted on instructional mediation to ensure that feedback

ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

was not meaningless. This implies that NLP tools can be best applied as an addition to teacher expertise and not as a replacement. The overlap between the views of the teacher and the student supports the dual requirement of clarity in the feedback and interpretation guidance.

The synthesized results illustrate that there is an intricate association between human perceptions and computational analysis. Coh-Metrix is efficient in bringing out cohesion and syntactic complexity patterns, but the users-teachers and students need a guide to interpret these results. The instantaneous nature of NLP feedback is appreciated by the students but their effects are curtailed by confusion and partial mistrust. The teachers like its ability to diagnose, but warn of over-reliance and the need to mediate with humans. All of these findings are indicative that, to the extent that NLP tools are well considered to complement pedagogical practices fostering interpretation, reflection, and critical engagement, they can be successfully used in the process of writing instruction.

Conclusion

The results of this research indicate that NLP generated feedback through Coh-Metrix gives meaningful information on the writing of students especially in cohesion and syntactic complexity. Nevertheless, the computational analysis can point to the presence of crucial linguistic patterns but students tend to have difficulties with interpretation without the help. Some liked the immediacy and clarity of automated responses and stated that they were confused by technical lingo and were not sure of accuracy.

This is an indication that NLP devices, as though helpful, cannot serve as independent assessors; students need teacher feedback to convert the computer-generated feedback into substantial corrections. Teachers also reported positive effects of Coh-Metrix as a diagnostic tool but its limitation in the evaluation of the higher-order writing skills like argumentation and rhetore structure. Their worries regarding overreliance of the students also highlight the fact that automated tools should be incorporated in a moderate pedagogical context. All in all, the research indicates that the application of NLP tools can help to improve the writing teaching process when it is supported with teacher mediation, which supports the idea that careful and informed application of such tools is necessary instead of blindly using automated assessment.

References

Bai, L., & Hu, G. (2021). Learners' engagement with automated feedback in L2 writing. *Assessing Writing*, *51*, 100934.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Crossley, S. A., Kyle, K., & McNamara, D. S. (2018). The development and use of cohesive devices in L2 writing and their relation to writing quality. *Assessing Writing*, *38*, 140–149. https://doi.org/10.1016/j.asw.2018.02.001

Crossley, S. A., Roscoe, R., & McNamara, D. S. (2019). Automated writing evaluation and feedback: Multiple metrics and multiple feedback types. *Computers in Human Behavior*, *90*, 226–238. https://doi.org/10.1016/j.chb.2018.09.024

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 193–202. https://doi.org/10.3758/BF03195564

Hyland, F., & Hyland, K. (2006). Feedback on second language students' writing. *Language Teaching, 39*(2), 83–101. https://doi.org/10.1017/S0261444806003399
Lee, I., Mak, P., & Yuan, R. (2021). Students' perceptions of automated feedback in L2 writing. *Journal of Second Language Writing, 51*, 100525.
Li, J., Link, S., & Hegelheimer, V. (2019). Learners' uptake of automated feedback in L2 writing. *Computers in Human Behavior, 90*, 226–238. https://doi.org/10.1016/j.chb.2019.03.012
Link, S., Dursun, A., Karakaya, K., & Hegelheimer, V. (2020). Student–machine agreement and consistency in automated writing evaluation. *Computers & Education, 151*, 103899. https://doi.org/10.1016/j.compedu.2020.103899
Lu, X., Sun, Y., & Liang, J. (2021). Natural language processing tools for writing evaluation: A systematic review. *Computers in Human Behavior, 120*, 106645.
McNamara, D. S., Allen, L. K., & McCarthy, K. S. (2017). Reading and writing in the digital age: Empirical and computational approaches. *Cognitive Systems Research, 43*, 1–12. https://doi.org/10.1016/j.cogsys.2017.04.002
Ranalli, J., & Arrigoni, E. (2019). Evaluating automated written corrective feedback: Accuracy, alignment, and pedagogical implications. *Computers & Education, 140*, 103653. https://doi.org/10.1016/j.compedu.2019.103653
Saadia Khan, Syed Moin Shah Bukhari, & Syed Ali Zain ul Abideen Naqvi. (2025). Exploring metaphorical language and conceptual framing in British English using nltk. *Journal of Applied Linguistics and TESOL (JALT)*, *8*(3), 644-655. https://doi.org/10.63878/jalt1017
Stevenson, M., MacLeod, G., & Evans, M. (2023). Students' interpretation of automated feedback and its impact on revision. *Computers & Education, 194*, 104771.
Wilson, J., & Roscoe, R. (2020). Automated writing evaluation and teachers' perceptions of feedback quality. *Computers in Human Behavior, 110*, 106270. https://doi.org/10.1016/j.chb.2020.106270
Zhang, R. (2023). Learners' engagement with automated writing feedback: A qualitative perspective. *Assessing Writing, 57*, 100722. https://doi.org/10.1016/j.asw.2023.100722
Zhang, Z., & Hyland, K. (2022). Exploring stance and engagement in student writing with automated feedback tools. *Journal of Second Language Writing, 55*, 101000.