



Research Article

Effects of Multimedia Glosses and Output Activity on Ethiopian Secondary School EFL Students' Reading Comprehension and Perceptions

Hawa Mohammed Yimam^{1,2}, Dawit Amogne³, Neil J Anderson⁴

¹Wollo University, Dessie, Ethiopia;

²Woldia University, Woldia Ethiopia;

³Addis Ababa University, Addis Ababa, Ethiopia;

⁴Brigham Young University, Hawai'i, USA;

ARTICLE INFO

Received: 10 March 2026

Revised: 13 June 2026

Accepted: 25 August 2026

Available online: 15 September 2026

KEYWORDS

EFL reading comprehension,
Multimedia glosses,
Output production,
Perception,
Resource constrained context

ABSTRACT

Multimedia glossing in second/foreign language learning has been used to support reading comprehension. However, most research has treated them as receptive aids and little attention has been paid to their pedagogical value when systematically integrated with output-based reading activities. Furthermore, existing studies have been largely conducted in technologically advanced settings, leaving low-resource educational contexts underexplored. This study investigated the effects of Multimedia Glosses and Output Activity (MGOA) on Ethiopian secondary school students' reading comprehension and perceptions. Following a pilot study, mixed-methods embedded quasi-experimental design was employed. Two intact grade 11 classes ($N = 94$), representing a population of 194 students, were randomly selected and assigned to a MGOA based experimental group (EG) and a conventional instruction based controlled group (CG). Both groups participated in eight instructional sessions. Data were collected through pre- and post-tests, reflective journals and semi-structured interviews and analysed using multivariate and mixed-design statistical procedures alongside thematic analysis. The quantitative results indicate that the EG significantly outperformed the CG in productive and recognition-based reading comprehension measures, with moderate ($r = 0.50$) and large ($d = 1.0$) effect sizes, respectively. Qualitative evidence further supports these results, revealing that students perceived MGOA-based instruction as engaging, motivating, and supportive for improving comprehension of reading texts in English. By combining multimedia support with output-oriented tasks, the intervention provides an effective model for improving language learning outcomes. The findings offer important implications for language teachers, curriculum developers, and policymakers seeking to integrate digital innovations into reading instruction.

1. Introduction

Reading is widely recognised as a foundational skill for learning because it enables students to access and interpret disciplinary knowledge. The Ethiopian general education curriculum emphasises the development of reading skills in upper secondary schools (Ministry of Education, 2020), given that learning in the education system is widely dependent on written materials (Gensler, 2023). A significant aspect of reading is the ability to comprehend meaning from texts (Anderson, 2026; Grabe & Stoller, 2019), which is significantly associated with higher students' academic achievement (Esleta, Onquit, Balitaon, & Aguila, 2024; Fedá, Gerasimova, & Vasiliev, 2022; Stoller,

* Corresponding author. hawamohammed99.hm@gmail.com (Yimam H. M.)

DOI: <http://doi.org/10.20372/ajids.2026.2930>

Shin, Grabe, & Copnway, 2026). This suggests that students who have strong reading comprehension skills consolidate evidence from texts effectively and transfer learning across subjects.

Nevertheless, developing reading comprehension skills remains challenging for second/foreign language (L2) students, particularly in low-resource contexts. Many students struggle to interact with reading materials, even after years of reading instruction (Alvarez Marinelli, Boggild-Jones, Crawford, Dubeck, Jhingran, Lack, Mohohlwane, Oviedo Buitrago, Piper, & Saavedra, 2025). This challenge is evident in Ethiopian secondary schools, where many students often demonstrate limited reading outcomes (Minda, 2019; Taye & Teshome, 2025), which negatively affects students' success in all areas of the curriculum and undermines national educational goals (Feda et al., 2022; Ministry of Education, 2023b). Despite local research into various reading strategies (e.g., Gebremariam & Weldeyohannes, 2025; Mohammed, 2020; Wendaferew & Damtew, 2023), students reading comprehension difficulties remain prevalent.

This problem could be attributed to different factors. For example, limited vocabulary knowledge prevents students from allocating sufficient mental resources to high-level coherence formation by increasing extraneous cognitive load (Mihret & Joshi, 2025; Minda, 2019; Taye & Teshome, 2025). The dominance of traditional instructional practices also affects the teaching of reading in L2 classrooms (Alvarez Marinelli et al., 2025). Consequently, there is a need for pedagogical solutions that address comprehension impediments.

To scaffold reading comprehension, literatures establish the use of glosses which are concise explanations of unfamiliar words embedded within a text (Grabe & Stoller, 2019; Nation, 2000). Traditionally, glosses have been used as written definitions in the form of marginal notes or footnotes. With the emergence of technology, glosses have evolved into multimedia formats that combine textual definitions with auditory or pictorial elements. Multimedia glossing is theoretically anchored in the input hypothesis (Krashen, 1985), the noticing hypothesis (Schmidt, 1990), the cognitive theory of multimedia learning (Mayer, 2014) and the dual coding theory (Paivio, 1990). Glosses make input understandable by eliminating lexical barriers (Krashen, 1985), draw learners' attention towards target input through salient features (Schmidt, 1990), and facilitates better learning and optimize cognitive processing through verbal and visual channels (Mayer, 2014; Paivio, 1990). These frameworks position multimedia glosses as effective receptive tools for comprehending texts.

Additionally, the role of multimedia glosses in facilitating text comprehension has been established by empirical studies (Khezlrou, Ellis, & Sadeghi, 2017; Rashid, Nisar, & Faiz, 2024; Sadeghi, Khezlrou, & Modirkhameneh, 2017; Tytko, Hui, & Pandža, 2026; Wang, 2020). For example, Tytko et al. (2026) reported that images and images with audio glosses significantly improved the reading comprehension of 65 students compared with no gloss groups. Rashid et al. (2024) also reported significant comprehension outcomes from combining audio with video glosses, followed by audio and picture glosses among 10 students. Another significant comprehension gain was reported by Wang (2020) in utilizing text, text plus audio, and text plus video glosses with 77 students. However, these studies were conducted in a foreign context with university level students. In Ethiopia, multimedia supported reading pedagogy is at an early stage. Although Tsegaw, Filate and Kahsay (2024) reported that glosses effectively mediate reading comprehension and vocabulary learning among 69 first year students, the intervention examined only one type of glossing.

Despite the benefits, glossed input alone could not guarantee robust comprehension, particularly where exposure to a target language is limited within classroom settings. Thus, students should be challenged to work with the input and retrieve it through subsequent output-oriented activities. The pedagogical values of output activities have been emphasised by increasing body of literature (Ellis, 2015; Fiorella & Mayer, 2022; Long, 1996; Swain, 1985). The output hypothesis (Swain, 1985, 1995) contend that producing language drives learners from semantic to syntactic processing, enabling them to recycle textual information and consolidate meaning appropriate to a context. According to the generative learning theory, comprehension is maximized when students actively generate relations among text information through output strategies like summarizing (Fiorella & Mayer, 2016, 2022). Empirical evidence further confirms that producing text summaries strengthen reading comprehension (Bogale, 2018; Budianto, Setyosari, Kuswandi, & Ulfa, 2022; Hukom & Ferdinandus, 2024; Kesevan, Ridzwan, Gopal, & Swaran Singh, 2023). However, the application of such output activities in multimedia learning environment is limited.

Consequently, the literature reveals three major limitations. First, although multimedia glosses have been widely studied as receptive tools, there is a lack of understanding regarding how they benefit comprehension when

explicitly embedded within output-based reading activities. Second, most multimedia gloss studies have been conducted in technologically advanced settings using web-based platforms and reading software; thus, little is known about how multimedia glosses benefit students in resource-constrained contexts where the digital divide necessitates offline digital solutions. Third, the majority of research focuses on tertiary settings. To our knowledge, reading instruction that synergistically integrate multimedia gloss and output-oriented activities remains scarce. Furthermore, understanding student perceptions is essential in designing instruction that fosters deep text comprehension.

Addresses these gaps, the current study examines how integrating multimedia glosses and output activity (MGOA) affects secondary school students' reading comprehension in Ethiopia, and how students perceive this instruction. MGOA represents the use of computer mediated, hyperlinked, offline glosses (textual, pictorial, or textual + pictorial glosses) during on screen reading and a subsequent output activity requiring students to generate a coherent reading summary incorporating target vocabulary items. The findings are likely to contribute to the literature by extending current discussions on L2 reading and multimedia learning theory. By combining multimedia support with output-oriented tasks, the study provides an effective model for improving language learning outcomes. The findings also offer important implications for language teachers, curriculum developers, and policymakers seeking to integrate context-appropriate digital innovations into reading instruction

2. Research Questions

The present study seeks to answer the following research questions:

1. Does instruction employing MGOA in a computer mediated reading result in greater gains in students' reading comprehension than a conventional instruction?
2. How do students perceive the use of MGOA for learning reading comprehension?

3. Methods

3.1. Design of the Study

The study employed mixed-methods embedded quasi-experimental design to examine the effects of computer-mediated MGOA instruction on reading comprehension and to explore students' perceptions of the intervention. The EG participated in a computer-mediated reading with multimedia glosses and a reading summary that required recycling of the glosses. The CG read the same text in print without glosses and output activity (see Fig. 1). Quantitative data were collected before and after the intervention, whereas qualitative data were collected during and after the intervention. The study was conducted from April to June 2024/2025.

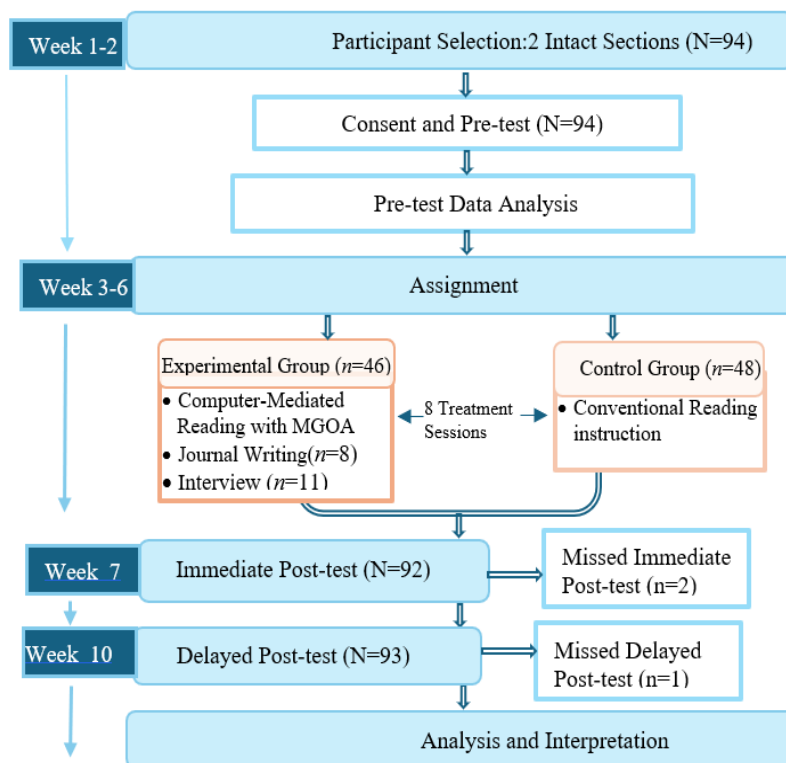


Figure 1. Research design.

3.2. Research Setting and Participants

The study was carried out at Woldia Higher Education Preparatory and Secondary School in Woldia city. The school was purposely selected because it had a sufficient number of computers, unlike the other three schools in the city. The participants were two intact Grade 11 sections randomly selected from four classes with a total population of 194 students. After reading comprehension baseline equivalence was established through pre-test assessment, the sections were assigned to an EG (n=46) and a CG (n=48) via a lottery system. All the students in the EG were asked to keep a journal, but only 8 representative entries were selected for the analysis. For the interviews, 11 participants from the EG were randomly selected. All the participants were between 16 and 21 years old. Because the study required computer literacy, Grade 11 students rather than Grade 9 or 10 students were purposefully selected to ensure that the participants could navigate computers effectively. Additionally, Grade 11 students were more accessible than Grade 12 students, who were preoccupied with preparations for the national examination. It is worth mentioning that three participants were excluded from the final analyses due to missing post-test data (two from the CG and one from the EG).

3.3. Materials and Instrumentation

Texts

Five reading passages, with word counts ranging from 397-558, were used for the intervention. The passages were standardised through rigorous steps. First, from the students' English Textbooks (Ministry of Education, 2023a), we randomly selected three expository passages: Features of Contemporary Civilization (p. 42), Indigenous Forest Conservation (p. 99), and The Effects of Global Warming on Economic Development (p.149). They hold a range of 364 to 578-words count (see Table 1). The passages were subjected to the Online Graded Text Editor (OGTE) to determine their difficulty level and lexical coverage. Since the OGTE grades texts according to specific language proficiency levels, we chose the Common European Framework of Reference (CEFR) for language teaching, learning, and testing because of its standardised use in describing language ability across various languages (Council of Europe, 2001). The framework includes six levels of language mastery: beginner (A1), elementary (A2), intermediate (B1), upper intermediate (B2), advanced (C1), and proficiency (C2). However, as the participants in

this study were Grade 11 students, only the upper four levels of the CEFR were considered relevant to the study. Second, using the results as a frame of reference, we prepared five reading passages, namely, 1) Renewable Energy for a Sustainable Future, 2) Mastering Time Management, 3) Understanding the Psychological Impact of Video Games on Youth, 4) Causes and Effects of Procrastination in Students, and 5) The Impact of Technology on Modern Education, with word counts of 397-558, with the assistance of ChatGPT-4.

We then systematically analysed each passage in the OGTE and modified the lexical difficulty to closely match the reference passages (see Table 1). Third, all the passages were evaluated by two Grade 11 EFL teachers, and a grant was obtained regarding linguistic difficulties and suitability for the students. Finally, they were pilot tested with a similar group of students (N=79) at a different school.

Table 1. Comprehensibility and Lexical Coverage of the Reading Materials (in %).

Reading Passages	Reference Passages					Intervention Materials				
	B1	B2	C1	C2	Word count	B1	B2	C1	C2	Word count
1	79.51	85.79	94.32	95.36	364	75.60	88.30	93.46	95.24	505
2	73.01	87.86	90.42	91.50	578	75.93	88.83	91.07	92.06	397
3	76.10	88.83	91.69	94.81	379	77.78	89.96	92.52	93.38	457
4						75.20	88.60	91.00	92.80	493
5						79.08	89.64	91.83	92.43	502

3.4. Target words for glossing

A total of 123 vocabulary items were selected for glossing. Initially, 204 out-of-level words for B1-level students, reflecting a word knowledge gap with vocabulary coverage of approximately 75% to 79%, were selected. It was assumed that glossing these words would increase the expected level of lexical coverage (95%) and improve text comprehensibility (Grabe & Yamashita, 2022; Nation, 2000). However, only 153 words were reserved, as 51 words with duplicate occurrences, closely related concepts, and common nouns were systematically removed. These words were presented to 79 similar grade-level students via a 3-level vocabulary knowledge scale adapted from Paribakht & Wesche, 1997. The students were asked to read a list of 153 words and rate the level that best described them: not knowing a word (level I), recognising a word but not understanding its meaning (level II), remembering the meaning of a word and being able to employ a definition or synonym of the word (level III). On the basis of the responses, 123 words that were known by only 20% or less of the participants (Khezrlou, 2018; Nation, 2000) were identified for glossing.

Adobe Acrobat Pro was used to annotate target vocabulary items with offline, hyperlinked hover glosses. Given the feasibility of multimodal glosses in resource-limited contexts, we embedded text, picture, or text + picture English glosses in digital texts. The annotated files were then readily available in PDF format and uploaded to computers. To make the target words more noticeable, they are highlighted in light blue. When students click on a word, an explanation, definition, or synonym in textual, pictorial, or textual + pictorial glosses pops up. These appear on the left, right, or bottom margins of the passages to avoid any intrusion during on-screen reading.

3.5. Data Gathering Instruments

Tests, reflective journals and semi-structured interviews were used to collect the data. For both the EG and CG, we administered two complementary reading comprehension tests, a recognition-based reading comprehension multiple choice test (MCT) and a production-based reading summary test (RST). The MCT measured recognition-based comprehension, whereas the RST assessed deeper productive comprehension. The tests were conducted as pre-test, immediate post-test, and delayed post-test. The pre-test was used to evaluate group homogeneity before the intervention and, later, to make inter- and intragroup comparisons with the post-test. The immediate post-test was employed to evaluate the short-term impact of the intervention on students' reading comprehension, whereas the delayed post-test was administered to assess the relative long-term effect of the intervention over time.

3.6. Reading comprehension multiple choice test

We designed the reading-comprehension multiple choice test (MCT) on the basis of the reading assessment principles of Alderson (2000); Brown and Abeywickrama (2019). The purpose of the MCT was to assess students' ability to understand texts at both the literal and inferential levels via objective items. There were 15 questions in the test. Each question had 4 possible options with one option being the correct response. The test was allocated a period of 30 minutes. Scoring was based on a dichotomous system where the right response got one point while the wrong response scored zero point. The maximum marks were 15. Reliability was measured via Kuder–Richardson's and the results were acceptable: pre-test (0.78), immediate post-test (0.81) and delayed post-test (0.80).

3.7. Reading Summary Test

The RST was employed to evaluate students' productive reading comprehension ability. The students were provided with 35 minutes to write a summary of a text they were given to read. A three-level rating rubric was adapted from the literature to rate the summary (Alderson, 2000; Brown & Abeywickrama, 2019; Taylor, 2013). The summary was scored by two EFL teachers who were trained and familiar with the rubric. The highest possible score was 10 points. The agreement between the two raters was assessed using a two-way mixed-effects model intraclass correlation coefficient, and the results indicated good interrater reliability across tests: pre-test (0.75), immediate post-test (0.77) and delayed post-test (0.80). Since the purpose of the RST is to measure students' reading comprehension by assessing the depth of their text understanding, the quality and organisation of students' writing were not considered.

3.8. Journals and Interviews

Reflective journals and semi-structured interviews were employed to explore the EG students' perceptions of the intervention. The reflective journals were utilised to capture students' experiences, thoughts, and reactions to computer-mediated reading with the MGOA. This tool was especially helpful for learning things that might not be immediately apparent in tests. Students filled up notebooks with leading prompts intended to generate in-depth observations on their experiences during the intervention. Students were instructed to write in either English or their native Amharic language in order to communicate their emotions without being hindered by language difficulties because the data were used to assess how the students felt about the intervention. All the entries written in Amharic were subsequently transcribed and translated into English for analysis.

Furthermore, the students' journal reflection was supported by data from semi-structured interviews after the intervention had been completed. Semi-structured interviews were conducted to gain in-depth information on the feelings of the students on learning to read with MGOA. For the purpose of making the interviewees comprehensible, interviews were conducted in Amharic then translated to English. All the interviews were recorded and transcribed.

3.9. Intervention Procedures

One week after the baseline equivalence data were collected, the experiment was implemented in the school's computer laboratory over eight treatment periods, each lasting 40 minutes. As shown in Fig. 2, the intervention was structured into three sequential stages, and the details are as follows.

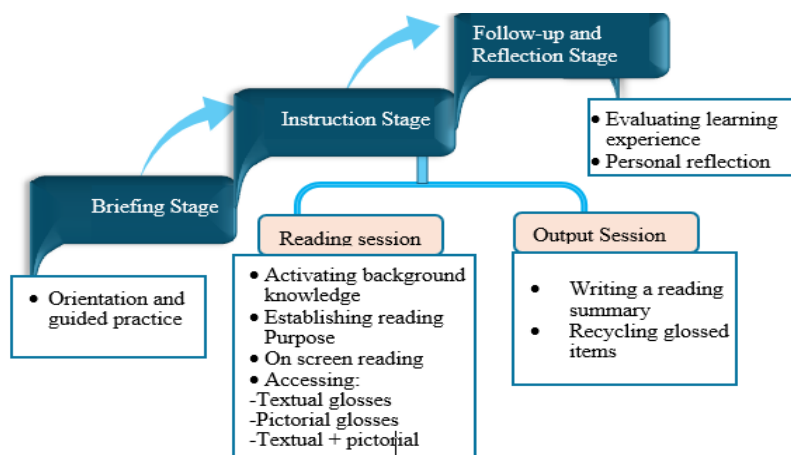


Figure 2. Intervention procedures

Briefing Stage: The briefing stage was implemented over two periods and consisted of an orientation session followed by guided practice. The primary objective of this phase was to establish an instructional context and familiarise students with the technological and pedagogical procedures of the intervention. Specifically, the students were trained on how to access multimedia glosses while reading texts on a computer screen and how to summarise reading while incorporating the target glossed lexical items. To scaffold the summarisation activity, a reading summary checklist was prepared on the basis of Brown and Abeywickrama's (2019) and Taylor's (2013) summarisation principles. It offered systematic guidance in order to help students find key points, organize concepts, and properly integrate glossed items.

Instruction Stage: In this stage, the instruction was provided into two sessions, i.e., a reading session and an output session, and it lasted for six periods. During the reading session, the students were presented with reading passages via desktop computers. To increase students' prior knowledge and encourage the formulation of reading, prereading questions were posed. Then, students proceeded to read the text and were instructed to click on unfamiliar lexical items to access additional information through textual glosses, pictorial glosses, or a combination of both to better understand the text. This was followed by the output activity session, where, students were told to use the glossed vocabulary words in a summary of the passage they read. To structure their responses, a summary-writing sheet with the target items was provided. The purpose of the output activity was to enhance overall comprehension through meaningful language production. At the conclusion of each session, students' written summaries were gathered, and formative feedback was given to encourage progress.

Follow-up and Reflection Stage: To evaluate the students' learning experience, reading comprehension activities were administered. Students were also asked to reflect on how they had learned reading comprehension through the use of computer-mediated MGOA instruction at the end of each output session. They were free to share their experience, viewpoints, challenges, and learning objectives through guided journal prompts. At the conclusion of the instruction session, these reflective notebooks were gathered.

One week after the completion of the intervention, the immediate post-test was administered. The delayed post-test was subsequently conducted two weeks later. For the CG, the same reading passages used with the EG were presented without glosses and the output activity. Instruction was delivered by the same teacher using a conventional teaching approach. After all the required data had been collected, the intervention was provided to the CG to ensure equitable access to the potential instructional benefits.

3.10. Methods of Data Analysis

The quantitative data were analysed through SPSS version 25. Before all the statistical analyses were computed, assumptions were tested. Assumptions of multivariate outliers, multicollinearity, homogeneity of variance and normality were all satisfied. Where violations occurred, appropriate measures were applied. Specifically, a violation of sphericity in the MCT post-test scores was addressed using the Greenhouse–Geisser correction. For the RST, as assumptions of normality and homogeneity of variances were violated, nonparametric tests were employed.

To examine baseline equivalence between the two groups, a one-way MANOVA was conducted. To assess the effects of the intervention over time, a mixed-design ANOVA followed by independent and paired-samples *t*-tests were performed. For the RST scores, Mann–Whitney U tests and Wilcoxon signed-rank tests were used. To account for possible Type I error in multiple comparisons, Bonferroni corrections were used. Each test effect sizes, i.e., partial eta squared (η^2) for ANOVA, Cohen's *d* for *t*-tests, and *r* for nonparametric tests, were calculated and interpreted as 0.2 for a small effect, 0.5 for a medium effect and 0.8 for a big effect size (Cohen, 2013).

Braun and Clarke (2022) thematic analysis guidelines were used to analyses the qualitative data. Initially, the two data sources were analyzed independently, and then similar codes were combined to create more comprehensive themes that represented students shared and varied experiences.

4. Results

4.1. Baseline Results

Table 2 presents the descriptive statistics of the EG and CG MCT and RST pre-test scores.

Table 2. Descriptive Statistics of the groups Pre-test Scores on the MCT and RST.

Measure	Groups	M	SD
MCT	Experimental (n=46)	6.02	2.09
	Control (n=48)	6.13	2.26
RST	Experimental (n=46)	3.98	1.70
	Control (n=48)	3.81	1.55

As shown in Table 2, the groups exhibited minor differences. For the MCT, the CG obtained a slightly higher mean score ($M = 6.13$, $SD = 2.26$) than the EG ($M = 6.02$, $SD = 2.09$). For the RST, the EG achieved a slightly higher mean score ($M = 3.98$, $SD = 1.70$) than the CG ($M = 3.81$, $SD = 1.55$).

A one-way MANOVA test indicated no statistically significant baseline difference in combined reading comprehension measures, Pillai's Trace = .004, $F(2, 91) = 0.20$, $p = .81$, $\eta^2 = .004$. This implies that the groups were comparable reading performance before the intervention. Thus, differences observed in the post intervention outcomes can reasonably be attributed to the effects of the intervention.

Effects of the MGOA on Reading Comprehension

Table 3 shows the descriptive statistics for the MCT scores by group and testing time.

Table 3. Descriptive Statistics of the MCT Scores by Group and Testing Time.

Measure	Group	Test Time	Mean	SD	N
MCT	Experimental	Pre-test	6.02	2.09	46
		Immediate post-test	8.98	2.894	46
		Delayed post-test	8.58	2.824	45
	Control	Pre-test	6.13	2.26	48
		Immediate post-test	6.22	2.624	46
		Delayed post-test	6.27	2.987	48

The results in Table 3 indicate that the EG achieved higher mean score ($M = 8.98$, $SD = 2.89$; $M = 8.58$, $SD = 2.82$) than the CG ($M = 6.22$, $SD = 2.62$; $M = 6.27$, $SD = 2.99$) at both the immediate and delayed post-tests, respectively.

For the MCT, a 2 (groups) $\times$ 3 (time) mixed-design ANOVA was conducted to assess changes across the three testing times and groups. The results revealed a significant main effect of time, $F(1.86, 165.68) = 53.07$, $p < .001$, $\eta^2 = 0.40$, and a significant main effect of group, $F(1, 89) = 10.35$, $p < .001$, $\eta^2 = 0.10$. Notably, a significant interaction between group and time was observed, $F(1.86, 165.68) = 43.36$, $p < .001$, $\eta^2 = 0.32$, indicating that the pattern of change over time differed between the groups.

A follow-up independent-samples *t* tests results showed that the EG significantly outperformed the CG on the immediate post-test, $t(90) = 4.79$, $p < .001$, $d = 1.0$ with a mean difference of 2.76, and on the delayed post-test, $t(91) = 3.82$, $p < .001$, $d = 0.79$ with a mean difference of 0.30.

Paired-samples *t* tests adjusted with Bonferroni correction were computed to compare MCT scores across the three time points by groups (Fig. 3).

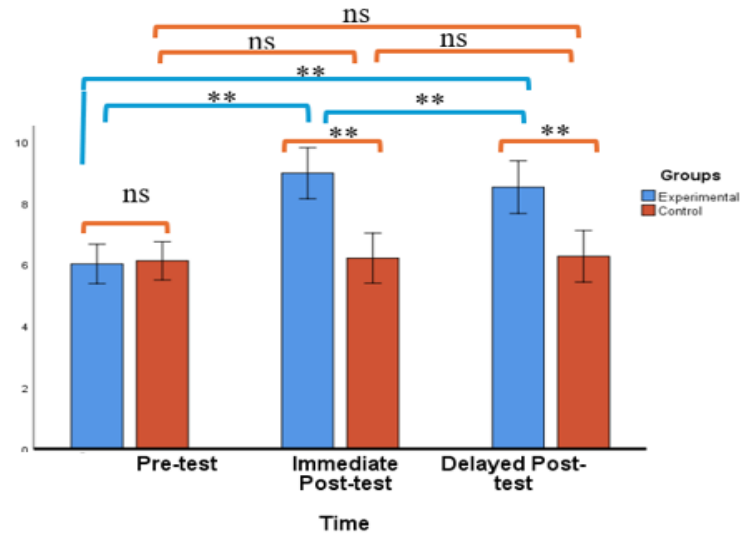


Figure 3. Results of the MCT across three time points by groups.

** indicates significance at $p < .001$. ns indicates nonsignificant value ($p > .017$). The error bars represent the 95% CIs. Significance was determined via the Bonferroni-corrected alpha level ($0.05/3 = 0.0167$).

The bar graph in Fig. 3 shows significant improvement in the EG from pre-test to both immediate post-test, $t(45) = 15.43$, $p < .001$, $d = 1.17$, and delayed post-test, $t(44) = 11.82$, $p < .001$, $d = 1.01$ with no significance difference between the immediate and delayed post-test scores, $t(44) = 2.00$, $p > .017$, $d = 0.17$. Whereas, for the CG, no significant change was found from pre-test to immediate post-test, $t(47) = 0.42$, $p > .017$, $d = 0.04$ and delayed post-test, $t(47) = 0.59$, $p > .05$, $d = 0.09$; or the immediate and delayed post-test scores, $t(47) = 0.87$, $p > .017$, $d = 0.05$.

For the RST, Mann–Whitney *U* test was computed to evaluate differences between groups. Table 4 presents the descriptive statistics.

Table 4. Descriptive Statistics of the RST Scores by Group and Testing Time.

Measures	Time	Group	Median	Mean Rank	Sum of Ranks
RST	Immediate post-test	Experimental (n=46)	6.00	58.32	2682.5
		Control (n=46)	4.00	34.68	1595.5
	Delayed post-test	Experimental (n=46)	6.00	59.28	2723.50
		Control (n=47)	4.00	35.49	1647.50

As displayed in Table 4, the EG outperformed the CG on both immediate post-test, $U = 514.50$, $z = -4.28$, $p < .001$, $r = .50$, and delayed post-test, $U = 527.50$, $z = -4.29$, $p < .001$, $r = .45$.

A Wilcoxon signed-rank tests with Bonferroni correction were computed to compare the tests across three time points between the groups. Fig. 4 presents the results.

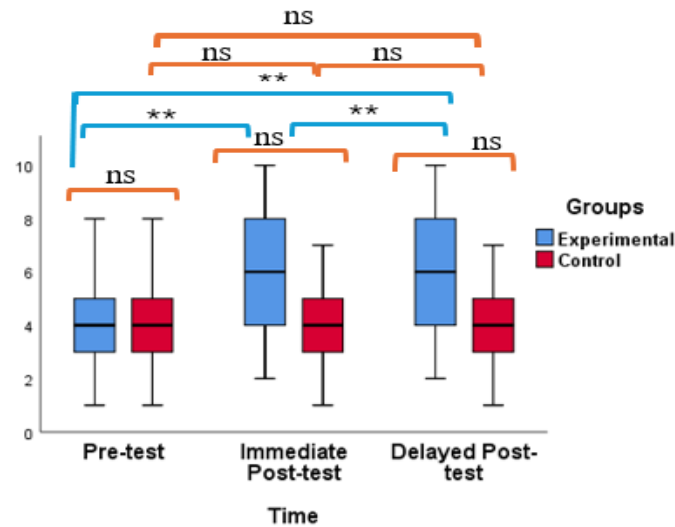


Figure 4. Results of the RST over time.

** indicates significance at $p < .001$. ns indicates nonsignificant value ($p > .017$). The error bars represent the 95% CIs. Significance was determined via the Bonferroni-corrected alpha level ($0.05/3 = 0.0167$).

The results in Fig. 4 revealed that the EG significantly improved from pre-test to immediate post-test ($Z = -5.34$, $p < .001$, $r = 0.6$, and delayed post-test ($Z = -5.56$, $p < .001$, $r = 0.5$). However, no significant difference was observed between the immediate post-test and delayed post-test ($Z = -.32$, $p > .017$, $r = 0.05$). For the CG, the differences from pre-test to immediate post-test ($Z = -.45$, $p > .017$, $r = .07$) and delayed post-test ($Z = -.92$, $p > .017$, $r = .08$) were not significant.

Overall, the quantitative findings indicate that computer-mediated reading with MGOA significantly improved students' reading comprehension across both recognition-based and production-based measures.

4.2. Students' Perceptions towards MGOA Instruction

The qualitative findings from reflective journals and semi-structured interviews indicated a predominantly positive perception characterized by the following four interconnected themes.

Perceived comprehension gains

Students consistently reported improved understanding of texts when using MGOA in a computer mediated reading. For instance, a student commented that "While the glosses supported my understanding of the texts, the summary encouraged me to read deeply to find out main ideas" (R3). Several students also commented on the output activity showing them the gaps in their initial level of understanding and encouraging them to re-read the text and look up the glosses. One student commented, "The output activity helps to compare what I read with what I understand, and it tests how many of the words I remember" (R9). The integration of multimedia glosses with the output activity helps students assess their reading progress and facilitate their reading process.

Vocabulary learning

The integration of multimedia glosses with the reading summary was perceived as an important strategy for learning new words and expanding vocabulary. The availability of text and pictorial information for unfamiliar vocabulary items was also seen to improve memorability. This was captured by a student who noted, "When I clicked on a word that I do not know, I saw its picture and definition. This made the word memorable" (R4). Students highlighted that using glossed vocabulary in summary writing helped move beyond recognition to productive use. As one participant stated, "when I use new words in the summary, I can remember them" (R9). This indicates that using target words in the summary of the reading text has influenced the use of multimedia glosses and transform them from receptive

aids to active lexical resources. This helps to enhance both receptive and productive vocabulary development through reading.

Reading motivation and engagement

Almost all students expressed excitement about computer-mediated reading with multimedia glosses and reported that the output activity basically changed how the students interacted with multimedia glosses. Students indicated that multimedia glosses made the reading process interactive, modern, and enjoyable. As one student stated, “Reading with multimedia gloss is very interesting” (R5). Another student further explained, “I was afraid to read before, but now I am interested” (R3). In addition, the availability of immediate lexical support fostered a sense of autonomy, as students could independently access meanings without relying on the teacher. As one participant explained, “I do not stack or wait for the teacher because the glosses help me understand by myself” (R9). This indicates that MGOA not only supported comprehension but also promoted learner independence and sustained engagement.

Overall evaluation

Overall, students viewed MGOA as an innovative and meaningful instructional approach. They expressed a preference for its continued use and recognized its value in improving their reading practices. All students recommended the integration of computer-mediated gloss-based reading with output activities in EFL classrooms. This positive evaluation reinforces the potential of MGOA as a pedagogically feasible strategy for enhancing reading comprehension in similar contexts.

5. Discussion

The present study demonstrates that computer-mediated reading with MGOA significantly enhances Grade 11 students’ reading comprehension. Students who participated in the experiment outperformed those who were in the CG on both MCT and RST immediate post-tests, with large ($d=0.90$) and moderate ($r=0.50$) effect sizes, respectively. The findings also demonstrated sustained gains across MCT and RST delayed post-tests ($d = 0.80$ and $r = 0.50$). This confirms MGOA enhanced both recognition-based and output-based comprehension, but it had a stronger effect on MCT than on RST. The benefit of MGOA was more readily captured by recognition-based tests than output-based tests. This may be due to the greater cognitive demands and the emphasis on identifying main arguments associated with summarization.

Similarly, the EG’s reading comprehension showed significant increment from the pre-test to both immediate and delayed post-test. The study revealed large effect sizes on the MCT immediate ($d = 1.17$) and delayed ($d = 1.00$) post-test, and moderate effect sizes on the RST immediate ($r = 0.60$) and delayed post-tests ($r = 0.50$). This indicates that MGOA has substantial short-term gains as well as sustained progress in students’ reading comprehension across different comprehension measures. In contrast, the CG did not show significant improvement from their initial pre-test to post-test on either test. The effect sizes on the MCT immediate ($d = 0.03$) and delayed ($d = 0.05$) post-tests, and the RST immediate ($r = .07$) and delayed ($r = .08$) post-test showed that the conventional instruction for teaching reading had little substantial impact on students’ improvement in reading comprehension.

The results of the qualitative data provide support for the quantitative findings revealing that students perceived MGOA-based instruction as engaging, motivating, and supportive for improving comprehension of reading texts in English. This suggests that the affective and cognitive engagement of MGOA facilitated improvement in reading comprehension.

The multimedia glosses provided comprehensible input (Krashen, 1985) which promoted students to process and evaluate the input, and reuse it through the output activity. This aligns with the noticing hypothesis (Schmidt, 1990) and the pushed output hypothesis (Swain, 1985, 1995) regarding the importance of output to obtain students’ attention for searching for meaning and evaluating context. Ultimately, learning occurs as a result of attention to input, processing of it, and storage in long-term memory (Anderson, 2005; Leow, 2024; Selinker & Gass, 2008). In addition, the findings support the interaction hypothesis (Chapelle, 1998; Ellis, 2015; Long, 1996), which posits that students’ interaction with input, such as negotiation of meaning, reflection, and integration of input, is an effective way of improving comprehension.

The results of this study also support previous studies (e.g., Khezlrou, 2018; Khezlrou et al., 2017; Sadeghi et al., 2017; Tytko et al., 2026), which highlighted the positive effects of multimedia glosses in reading comprehension. While these studies were conducted in technology-rich university contexts, the present study demonstrates that offline, computer-mediated MGOA can be effective even in resource-limited secondary school settings. Furthermore, students' positive perceptions of gloss-based instruction observed in this study aligns with Khezlrou (2018); Wang (2020); Wang and Lee (2021), who established dual benefits of glosses in enhancing comprehension and being favourably received by learners.

General, the study suggests that the integration of multimedia glosses and output-based reading activities changes reading from a passive recognition activity to an interactive and productive process in which students monitor their understanding, make use of new vocabulary, and cognitively and affectively engage with what they read. This suggests that MGOA is a viable and sustainable approach for improving reading comprehension and students' engagement in EFL classrooms.

6. Conclusion

This study examined the effects of MGOA on secondary school EFL students' reading comprehension and perceptions. The findings indicate that MGOA significantly improves reading comprehension, and it is positively perceived by students as they feel improvements in their reading comprehension, vocabulary learning, reading motivation and engagement. Overall, the results highlight the primacy of MGOA instruction over the conventional method of teaching reading.

The findings of the study contribute to the literature by demonstrating that integrating multimedia glosses explicitly with output activity provides an effective model for improving language learning outcomes in resource-limited educational contexts. The findings benefit L2 learning through multimedia research by reconceptualising glosses as integral to an input–output learning cycle rather than a receptive aid per se. It also provides insights into the preparation of offline multimedia glosses in technologically under resourced educational settings, where digital learning management systems and adaptive reading software are not affordable.

Therefore, MGOA-based reading instruction can be applicable to comparable EFL contexts that share similar sociolinguistic and institutional conditions. EFL teachers are encouraged to move beyond traditional reading instruction and prepare computer-mediated reading lessons encompassing hyperlinked textual, pictorial, or textual + pictorial glosses and integrate output-based reading activities to draw students' attention to main ideas and overall messages. Thus, targeted teacher training is necessary. Curriculum designers are urged to design e-textbooks that embed multimedia-supported reading materials integrated with output activities for comprehensive and sustainable improvement in students' reading comprehension.

7. Limitation and Future Directions

Despite the significant contributions of MGOA instruction to improve secondary school students' reading comprehension, our findings should be interpreted cautiously due to certain limitations. The major limitation of this research is that it only involved a single school and only two groups. Therefore, in order to improve the generalization of the research to similar educational settings, further research could be carried out in different schools and involving a greater number of participants using a similar procedure. A more thorough knowledge of the intervention would also result from broadening the sample to address individual difference and include diverse students from various geographical areas such as urban and rural.

Another key limitation of the study was that we examined textual, pictorial, or both textual and pictorial glosses in a single experimental group. Although the integrated use of these glosses supported students reading comprehension, the relative effectiveness of each gloss type in an offline digital context remains questionable, and which gloss type works best with output activities also needs further consideration. Future studies could use a more analytically design that methodically links each gloss types with output activities beyond integrated gloss conditions.

The third limitation of the study was related to glossing formats. The introduction of audio glosses was limited by logistical and resource-related issues, although audio glossing has the potential to benefit auditory learners and enhance students' multimodal engagement. In particular, using audio glosses necessitates having access to sufficient

technology, such as headsets or a language lab, which is not possible for every participant in this situation. As a result, the study was restricted to non-audio glossing modalities. Thus, to evaluate the additional instructional usefulness of audio glosses, future research is recommended.

The other limitation was that the study was conducted within a short period of time. In order to examine the long-term impacts of the intervention, longitudinal designs incorporating follow-up assessments, classroom observations, and evaluations of students' performance across different language skills could be used in future studies. Additionally, further research could be conducted to examine how the learned comprehension skills can be applied to other language skills, such as writing.

Despite the limitations, this study provides a valuable approach for teaching reading comprehension in secondary school EFL classrooms, particularly where technological resources are scarce.

Acknowledgements

We acknowledge the TIRF-British Council for providing partial funding for this study.

Data availability statement: will be provided upon reasonable request.

Declarations

Ethics Approval: This study was reviewed and approved by the Institutional Review Board of Woldia University under the protocol number WDU/IRB027/01/2025. All methods of the study and procedures concerning human participants were performed in accordance with the standards and ethical guidelines of the Woldia University Institutional Review Board, which complies with the national research and ethical guidelines.

Credit Author Statement: Hawa M. Yimam and Dawit Amogne: Conceptualisation and Methodology. Hawa M. Yimam: Investigation, Analysis and Original Draft. Neil J Anderson: Resources. Dawit Amogne and Neil J Anderson: Reviewing and Editing

Consent to Participate: Before the study, informed written consent to participate was obtained from all participants and parents or legal guardians of those under 18.

Consent to Publish: Informed written consent to publish anonymized data was obtained from all participants and parents or guardians of those under 18 years of age.

Conflict of interest: The authors declare no conflicts of interest.

Reference

- Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511732935>
- Alvarez Marinelli, H., Boggild-Jones, I., Crawford, M., Dubeck, M. M., Jhingran, D., Lack, C. J., Mohohlwane, N., Oviedo Buitrago, M. E., Piper, B., & Saavedra, J. (2025). *Effective Reading Instruction in Low-and Middle-Income Countries: What the Evidence Shows*. Global Education Evidence Advisory Panel (GEEAP) / World Bank Group. <https://documents1.worldbank.org/curated/en/099448110272527300/pdf/IDU-740c63ca-259a-4fae-8d95-7935242a25db.pdf>
- Anderson, J. R. (2005). *Cognitive psychology and its implications*. Macmillan.
- Anderson, N. J. (2026). Developing engaged L2 readers. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (5th ed., pp. 148–165). Boston MA: National Geographic Learning/Cengage Learning.
- Bogale, Y. N. (2018). Conceptualizing reading to learn: strategy instruction and EFL students' reading comprehension. *International Journal of Curriculum and Instruction*, 10(2), 93–117.
- Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative psychology*, 9(1), 3.
- Brown, H. D., & Abeywickrama, P. (2019). Language assessment: principles and classroom practices. In: Pearson.
- Budianto, A., Setyosari, P., Kuswandi, D., & Ulfa, S. (2022). Summaries Writing to Enhance Reading Comprehension: Systematic Literature Review from 2014 to 2021. *Eurasian Journal of Applied Linguistics*, 8(1), 149–161.
- Chapelle, C. (1998). Research on the Use of Technology in TESOL. *RESEARCH ISSUES*, 32(4), 753.

- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. routledge. <https://doi.org/10.4324/9780203771587>
- Council of Europe. (2001). *Common European Framework of Reference (CEFR) for Languages: Learning, Teaching, Assessment*. Cambridge University Press.
- Ellis, R. (2015). *Understanding second language acquisition 2nd edition*. Oxford university press.
- Esleta, L., Onquit, I., Balitao, A., & Aguila, A. (2024). Reading ability and academic achievement [Conference session]. The Asian conference on education and international development 2024 official conference proceedings, Tokyo, Japan,
- Feda, K., Gerasimova, J., & Vasiliev, K. (2022). *Learning Poverty in Ethiopia : Status, Key Factors, and Priority Reduction Strategies (English)*. World Bank Group. <http://documents.worldbank.org/curated/en/099615106172213756>
- Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational psychology review*, 28(4), 717–741. <https://doi.org/10.1007/s10648-015-9348-9>
- Fiorella, L., & Mayer, R. E. (2022). The Generative Activity Principle in Multimedia Learning. In L. Fiorella & R. E. Mayer (Eds.), *The Cambridge Handbook of Multimedia Learning* (pp. 339–354). Cambridge University Press.
- Gebremariam, H. T., & Weldeyohannes, M. Z. (2025). Cultivating Students' Reading Skills: Inspiring Reading Comprehension and Motivation Through Critical Reading Strategies. *Sage Open*, 15(3), 21582440251367189.
- Gensler, O. (2023). Ethiopian English. . In R. Meyer, W. Bedilu, & L. Zelealem (Eds.), *The Oxford Handbook of Ethiopian Languages*. . Oxford University Press.
- Grabe, W., & Stoller, F. L. (2019). *Teaching and researching reading*. Routledge.
- Grabe, W., & Yamashita, J. (2022). *Reading in a second language: moving from theory to practice*.
- Hukom, S. J., & Ferdinandus, M. S. (2024). THE EFFECTIVENESS OF IMPLEMENTING SUMMARY WRITING TO IMPROVE STUDENTS' READING COMPREHENSION. *PEDAGOGIKA: Jurnal Pedagogik dan Dinamika Pendidikan*.
- Kesevan, H. V., Ridzwan, N. D. R., Gopal, R., & Swaran Singh, C. K. (2023). Written retelling strategy in improving reading skill among ESL students in Malaysia. *Education Research International*, 2023(1), 6694646. <https://doi.org/10.1155/2023/6694646>
- Khezlrou, S. (2018). Form-focussed instruction in CALL: What do learners think? *Relc Journal*, 50(2), 235–251. <https://doi.org/10.1177/0033688217738820>
- Khezlrou, S., Ellis, R., & Sadeghi, K. (2017). Effects of computer-assisted glosses on EFL learners' vocabulary acquisition and reading comprehension in three learning conditions. *System*, 65, 104–116.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. . Longman.
- Leow, R. P. (2024). Depth of processing in L2 learning. *The TESOL*. <https://doi.org/DOI:10.1002/9781118784235.eelt0978.pub2>
- Long, M. (1996). *The role of the linguistic environment in second language acquisition*. New York: Academic Press.
- Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning*. Cambridge. CUP.
- Mihret, G., & Joshi, J. (2025). Understanding and addressing reading comprehension challenges: Magnitude, causes and solutions. *Journal of English as A Foreign Language Teaching and Research*, 5(1), 1–11.
- Minda, M. H. (2019). An analysis of reading strategies used by Ethiopian higher education students: Evidence from Ambo University, Oromia, Ethiopia.
- Ministry of Education. (2020). *General Education Curriculum Framework*. Addis Ababa Retrieved from <https://edumediadepot.gei.de/server/api/core/bitstreams/3a21f225-3730-4f7f-9d32-44b81cc7ef0b/content>
- Ministry of Education. (2023a). *English Student Textbook Grade 11*. Gravity Group IND LLC.
- Ministry of Education. (2023b). *Ethiopia Education Transformation Programme (EETP): Strategic implementation framework*. Federal Democratic Republic of Ethiopia Retrieved from <https://www.dpgethiopia.org/wp-content/uploads/2024/03/Ethiopia-Education-Transformation-Program-proposal-.pdf>
- Mohammed, H. (2020). The effect of discourse analysis informed instruction on developing high school learners reading comprehension. *International Journal of Secondary Education*, 8(2), 20–26.
- Nation, I. S. (2000). *Learning vocabulary in another language*. Cambridge university press.
- Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford university press.
- Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. *Second language vocabulary acquisition: A rationale for pedagogy*, 55(4), 174–200.

- Rashid, K. U., Nisar, F., & Faiz, F. (2024). EFFECTIVENESS OF MULTIMEDIA GLOSSES IN READING COMPREHENSION OF ESL LEARNER AT THE UNIVERSITY OF AJ&K. *International Premier Journal of Languages & Literature*, 2(2), 222–235.
- Sadeghi, K., Khezrlou, S., & Modirkhameneh, S. (2017). Calling Iranian learners of L2 English: Effect of gloss type on lexical retention and reading performance under different learning conditions. *Journal of Research in Reading*, 40, S66–S86. <https://doi.org/https://doi.org/10.1111/1467-9817.12088>
- Schmidt, R. W. (1990). The role of consciousness in second language learning1. *Applied linguistics*, 11(2), 129–158.
- Selinker, L., & Gass, S. M. (2008). Second language acquisition. *Lawrence Erlbaum Ass.*
- Stoller, F. L., Shin, J., Grabe, W., & Copnway, J. (2026). Teching reading for academic purposes. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), (5th ed., pp. 166–188). Boston, MA: National Geographic Learning/Cengage Learning.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. *Input in second language acquisition*, 15, 165–179.
- Swain, M. (1995). Three functions of output in second language learning. In G. S. Cook, Barbara (Ed.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford University Press.
- Taye, T., & Teshome, G. (2025). Assessing Secondary School Students' Awareness and Challenges in Using Intensive and Extensive Reading Methods in Ethiopia.
- Taylor, L. B. (2013). *Testing reading through summary: Investigating summary completion tasks for assessing reading comprehension ability*. Cambridge University Press.
- Tsegaw, S. A., Filate, A. Y., & Kahsay, M. T. (2024). Effects of tool mediation on tertiary level EFL students' reading comprehension and vocabulary learning skills: a case for a cloud computing environment. *Cogent Education*, 11(1), 2330251.
- Tytco, T., Hui, B., & Pandža, N. B. (2026). The effect of multimodal input on L2 learners' reading comprehension: A preregistered eye-tracking study. *The modern language journal*, 110(1), 210–239. <https://doi.org/10.1111/modl.70023>
- Wang, S. (2020). Exploring the effectiveness of hypermedia glosses for second language reading. *Journal of Asia TEFL*, 17(4), 1284–1293. <https://doi.org/10.18823/asiatefl.2020.17.4.8.1284>
- Wang, S., & Lee, C. I. (2021). Multimedia gloss presentation: Learners' preference and the effects on EFL vocabulary learning and reading comprehension. *Frontiers in Psychology*, 11, 602520. <https://doi.org/10.3389/fpsyg.2020.602520>
- Wendaferew, D., & Damtew, A. (2023). The effect of explicit strategy instruction on students' reading comprehension performance of Ethiopian secondary school students. *Education Research International*, 2023(1), 1894026.