

AI-Generated Visual art in English Language Teaching: a Multimodal Approach to Language Learning

G.M. Prem Kumar¹, Dr. N.Hema²

¹Research Scholar, Department of Language, Culture and Society, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu, Tamil Nadu, India

Email ID:- iamprem009@gmail.com , ORCID: 0009-0006-9723-9983

²Professor, Directorate of Learning and Development, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu, Tamil Nadu, India

Email ID:- heman@srmist.edu.in , ORCID: 0000-0003-3536-8676

ABSTRACT

The fast advancement of Generative Artificial Intelligence (GenAI) has opened up new ways of combining visual and linguistic modalities for English teaching. The focus of this research is to study how effective the AI-generated visual art can be as a multimodal instructional tool for enhancing English language skills (listening, speaking, reading, and writing) as well as multimodal literacy in secondary students especially. In line with the Multiliteracies Theory, Social Semiotic Theory of Multimodality and Constructivist Learning Theory, a quasi-experimental pre-test-post-test control group design was used by engaging 100 ESL learners at the Chennai university in India. The experimental and control groups comprised of 50 learners each. The learners in the experimental group were taught English through the use of AI-generated visual images for vocabulary acquisition, reading, writing, grammar, creative expression, and multimodal understanding practice, while the control group participants were taught in the traditional way. The research utilized an English language proficiency test, a set of tasks focused on four language skills assessment, multimodal literacy rubric, questionnaire about students' impressions, and observations for data collection. The quantitative data from this study were analyzed with the help of different statistical techniques, such as descriptive statistics, paired-sample and independent-sample t-tests, gain-score analysis and ANCOVA. The results showed significant differences between the achievement of the experimental and control groups, with the first group performing better, $t(98) = 8.60, p < .001, d = 1.72$. The differences in multimodal literacy were also significant, $t(98) = 7.03, p < .001, d = 1.41$. The results of ANCOVA showed that even after controlling for the pre-test results, there was a significant effect of the type of instructions on the achievement of the participants in the post-test, $F(1, 97) = 85.18, p < .001, \text{partial } \eta^2 = .468$. The students in the experimental group had positive feelings and perceived improvement in their creativity skills.

Keywords: Generative Artificial Intelligence, AI-generated visual art, English Language Teaching, multimodal literacy, LSRW skills, visual learning, language education

1. INTRODUCTION

Digital technology has radically changed the methods of human communication, learning, and sense-making. Literacy is no longer limited only to reading and scribbling; modern communication involves multiple modes such as linguistic, visual, spatial, auditory and digital modes. According to New London Group (1996), the pedagogical approach of multiliteracies should be used in education due to the general diversity of modes used in communication and the necessity for people to operate with several modes to convey the meaning.

The mentioned transformation results in important consequences for ELT. Students see English more and more in such contexts as images, videos, websites, social media, digital stories etc. The use of visual materials such as pictures, images, cartoons or sequences of pictures has always been effective in learning new words, speaking, comprehension and creativity. However, now Generative Artificial Intelligence gives an opportunity to create visual materials corresponding to specific learning goals.

Visual art generated by AI has the potential to convert a teacher's written prompt into images that can signify characters, locations, acts, emotions, or narrative situations. These images can be used as multimodal texts for learning a language. To illustrate, an AI picture of a populated railway station can be applied to learn vocabulary, describe, practice grammar, create dialogues, or write stories. One can also create a sequence of images and narrate a story with the help of these visual aids.

The recent investigations show a rising interest in using GenAI in language education, especially for such purposes as writing, getting feedback, language application, and teaching methods (Zhang et al., 2025; Deng & Jamaludin, 2026). Nevertheless, the pedagogical usage of AI-generated visual art in relation to improving multimodal literacy in English language classrooms has not been studied deeply. This provides a good ground to analyze how students will interpret, assess, develop, and change meanings in the process of communication by means of pictures.

Nonetheless, the incorporation of images made by AI needs to be scrutinized in a subtle manner. The systems powered by AI often generate culturally unacceptable or stereotypical images. In addition, excessive reliance on

artificial intelligence might put a strain on the creativity and agency of learners. Therefore, UNESCO (Miao & Holmes, 2023) places emphasis on the need for human-centered and many-sided utilization of Generative AI in the process of teaching.

In this regard, this article pursues an in-depth investigation of AI-generated images and their potential in the process of teaching English. The paper aims to analyze the contributions of AI-generated visuals to the acquiring of vocabulary and such skills as speaking, writing, and the ability to act creatively. The work considers how the teacher plays the role of mediator and what ethical dilemmas arise as a result of the use of AI-generated images. The paper argues that when applied properly, the AI-generated images can offer students plenty of possibilities and opportunities for interpreting, sharing, creating, and transforming information.

2. LITERATURE REVIEW

2.1 Understanding Multimodality in the Context of Language Teaching

Multimodality is a concept that has lately gained popularity in the realm of language education because communication is never limited to the written or spoken word only. New London Group (1996) introduced the idea of multiliteracies in the context of the rapidly changing communication technologies and increasing diversity in cultures and languages. According to them, many different modes can be used to make meaning in communication.

There are several studies dedicated to examining the application of multimodality in English classes. These studies indicate that multimodal pedagogy brings students into the process of learning through visual, linguistic, digital, and other modalities. Yi and Ang (2022) conducted a systematic literature review, collating research from 2010 to 2021, and found that the implementation of multimodal approaches in English classes is mostly aimed at promoting critical, creative, and culturally relevant learning. The researchers also noted that another important issue in the field is multimodality literacy itself.

2.2 Generative AI in Language Education

The technology-assisted language education concept received an upgrade thanks to Generative AI emergence. Systematic literature reviews done in recent years indicate this academic area is proliferating rapidly. For instance, Li et al. (2025) reviewed 144 peer-reviewed articles from 2023 to 2024 and indicated a few fields the research is undertaking, including language skill learning, teaching techniques applied, writing and feedback, students' attitude, and others. Their review showed, however, that the writing aspect of the topic is researched the most, without paying proper attention to speaking, listening, and reading skills, especially in K-12 education.

Supportively to Li et al. (2025), the results of Lee et al. (2025) review of 49 empirical studies based on educational resource use indicate that the education segment of higher education dominates the field. They also expressed their opinion on the necessity of undertaking the follow-up research on GenAI multimodal features and research shifting focus from writing to a broader range of language skills.

2.3 Generative AI and Multimodal EFL

Various setups attribute the emergence of multimodal capabilities of GenAI to the way they facilitate the incorporation of both text and different kinds of visual data in language training practice. For instance, GenAI-generated images provide instructors with opportunities to design unique settings, personalities, plots, and stories based on classroom objectives. Therefore, students will be able to interpret the images, communicate their ideas about them, and convert them into spoken or written English.

Recent findings also confirm that the efficiency of language training supported by GenAI depends on the instructional design and the level of AI literacy of learners. In the systematic review and meta-analysis on task-based language teaching conducted in 2025, it was noted that GenAI has a positive impact on language training, whereas the characteristics such as AI literacy and proper planning of tasks can serve as moderators influencing the effectiveness of the technique utilized.

2.4 Research Gap

The literature available discusses the significance of multimodal techniques in teaching and explains the use of Generative AI in language gaining. But there is still a research gap in this area. The research on Generative AI in language learning has mostly been concerning text-based resources like writing, feedback, and communication. Thus, the role played by the use of AI-assisted visual arts in education is not examined deeply enough. The research study presented aims to fill this gap by investigating the role of AI-created visual images in helping the students to interpret the visual information effectively and producing meaningful English language content.

3. THEORETICAL FRAMEWORK

The present study is based on the assumption that learning English is increasingly about making meaning with the help of various modalities. As this paper relates to the use of AI-generated artworks for teaching English, the

framework for this study relies on Multiliteracies Theory, Social Semiotics Theory of Multimodality, and Constructivist Learning Theory. These theories give meaning to the study as they suggest how learners may use images created with the help of AI to make meaning in English.

3.1 Multiliteracies Theory

The main theoretical basis of this study comes from the Pedagogy of Multiliteracies of New London Group (1996). This framework appeared as a result of the awareness that communication is determined by cultural diversity and advances in communication technologies. It suggests that one should think of literacy as practically the whole range of modes of communicating instead of only relying on written form.

This viewpoint is especially important with regards to foreign language education, and particularly English, because students interact with English through text, pictures, sound, video and digital media. AI-generated images serve as a means for students to analyze pictures and turn them into their verbal and written forms in English. In this way, students learn to work not just as text readers but also deploy their skills in constructing messages in multimodality.

3.2 The Social Semiotic Approach to Multimodality

The research accordingly makes use of social semiotic concepts of multimodality, and particularly the theory advocated by Kress (2010). The theory considers language, visual texts, colour, gesture and space to be among semiotic means through which information is conveyed, each having unique possibilities for conveying meaning. Hence, an AI image can be thought of as a multimodal text and not only an illustration. Students can analyze pictures (images) and translate this information into English language. Working with visual and verbal systems can help students develop their vocabulary, descriptive skills, narrative skills and creativity in using the language.

3.3 Constructivist Learning Theory

The research also builds upon the Constructivist Learning Theory that sees learning as an active procedure in which learners build knowledge from experience, collaboration, and interpretation. The learners take part in the visual trigger, build the interpretation, engage in conversations, and create their own language as opposed to receiving the language knowledge without active participation.

For example, the students might look at an AI-generated image of some mysterious house, assess the vocabulary, assess what possibly happened in the scene, construct the story and propose a different ending to it. Hence, the image will serve as a starting point for a language learning journey.

3.4 Integrated Theoretical Approach

The theories work together to provide a full picture. The Multiliteracies Theory defines the importance of the use of multiple modes, the Social Semiotics Theory defines how meaning is formed through the work with the modes. The Constructivism Theory shows how learners construct knowledge by learning through the modes.

As a result, it can be said that through the presented framework artificial intelligence-generated art is displayed not only as a technological instrument but as a multimodal teaching resource that can be used by learners in the process of interpretation, negotiation, creation, and communication.

4. AI-Produced Visual Art for LSRW and Terms of Speech Progression

Implementing AI-produced visual art in English language instruction is a way to bridge Listening, Speaking, Reading, and Writing skills into one multimodal educational format. AI images and the sequences they are in can deliver a common environmental context for language students and allow them to get practice in the skills incorporated in the activities. Providing learners with visual stimuli promotes the linking between linguistic expressions and meaning in context and invokes the participation in tasks.

4.1 AI-Produced Visual Art and Listening Skills

AI visuals can aid listening activities due to the clues they provide learners with. For instance, a teacher may show learners an image or a series of images while reading out loud or clicking an audio of a description in English. In this case, learners will be doing activities like stating which of the images correspond to the description provided, identifying important pieces of information, or putting images in the correct order according to what they hear.

For instance, pupils may hear a narration of a busy railway station and recognize the people, things, and operations illustrated in the AI-generated image. Then the task can be extended with questions of comprehension and discussions.

4.2 AI-Generated Visual Art and Speaking

Visual prompts can give learners a good source for oral communication. For example, pupils may speak about the image they observe, talk about their impressions, guess what is happening in the image, or create a story according

to a collection of images. All these exercises promote vocabulary use and grammar practices in a more or less natural context.

To give an example, an image with a child standing in front of a strange house inspires a number of inquiries beneath:

Who is this person?

Where are they?

Why are they there?

What can happen next?

Using these questions, pupils are able to speak about their interpretation of the image in class, either individually or together.

4.3 AI-Generated Visuals and Reading

Visuals created through AI can be used in conjunction with reading activities. The teacher may show students a short passage in English and ask them to identify or interpret the visual representation of the text. Alternately, students may observe a piece of visual information and try to guess what a corresponding reading may be about. This activities allow students to make connections between the information presented visually and the one presented textually.

In addition, exploring both types of information can help to develop other strategies such as inferring, predicting, organizing information, and understanding the meaning.

4.4 AI-Generated Visuals and Writing

The most straightforward way of using AI-generated visuals is in the area of writing.

Visuals can help students generate ideas for doing description or creative writing. For instance, the students may write about people and places, construct stories on the basis of visual content in a sequence of images, write dialogues or diary entries based on a visual representation, or even write a storyboard.

4.5 Vocabulary and grammar development

AI-generated visual images can assist with vocabulary and grammar development through contextualisation at the time of teaching. Instead of teaching vocabulary in isolation, teachers can incorporate the visual images showing meaningful context to demonstrate vocabulary. For example, in an image illustrating a marketplace, teachers can present vocabulary on people, objects, actions, and places involved.

Visual images may also support grammar contextualisation. For instance, students can describe the actions presented in picture in tenses, identify location with prepositions, etc. Therefore, both grammar and vocabulary can be integrated into communicative activities rather than taught separately in isolation.

4.6 Integrated LSRW Learning

AI-generated visual art has great pedagogical power due to the possibility of incorporating the four language skills into a single lesson. Students will see a visual story, be given an audio description of it, speak about the images, read something connected to the story and create their own story.

Thus, AI-generated visual art can serve as a common impulse for integrated language learning. Its power is not limited to sheer provision of beautiful illustrations, but it creates conditions for language learners to relate the process of visual interpretation to meaningful English language production

5. MULTIMODAL LITERACY AND AI-BASED LANGUAGE LEARNING

Utilising an AI-based language learning strategy allows for implementing language and visual modes into the English language teaching process. The visual art produced by AI can serve as a multimodal text allowing students to interpret images and switch the visual data into verbal language forms. This approach helps pupils improve their connection between seeing, decoding, speaking, reading, and writing instead of studying these aspects of language separately.

As AI-generated images are influential in developing different kinds of literacy, students enhance their visual literacy by learning to interpret the visual as well as their linguistic literacy by communicating their interpretations. Additionally, pupils also develop digital and AI literacy through their critical engagement with the AI product. For instance, students can watch an AI scene, define and describe the participants and where it takes place to develop some hypotheses and write about their perceptions.

At the same time, it should be noted that the AI-based language learning process must be monitored by a teacher. AI abilities can sometimes lead to cultural stereotypes in generated images. That is why students need to be taught to evaluate and modify the information produced by AI and not just accept it as it is.

6. RESEARCH METHODOLOGY

6.1 Research Design

The study utilized a quasi-experimental mixed-method research design to investigate the role of AI-generated visual art in English Language Instruction. In terms of quantitative, the research assessed the level of the students' performance in English language competencies in addition to the multimodal literacy skills. As for the qualitative aspect, students' perceptions and experiences in class regarding the use of AI-generated visual art in the English language teaching process were involved in the analysis.

A pre-test and post-test design including experimental and control groups was put into practice in this research study. Specifically, the experimental group adopted English instruction based on the usage of AI-generated visual art, while the control group used traditional instructional materials in English language teaching.

6.2 Participants

The sample comprised 100 Tertiary students who were learning English as part of regular studies. The subjects of the study were divided equally into two groups, with 50 students in the experimental group and 50 in the control group.

All the subjects were students who were about 18 years old. The study was carried out in the city of Chennai.

The study adopted a convenience sampling technique. The study subjects were selected in such a way that their pre-intervention levels of English language proficiency were more or less comparable. Pretest was conducted to evaluate their initial performance

6.3 Research Context and Intervention

The basic teaching of English took place during this period involving the use of AI technologies. The experimental group utilized classroom activities with visual materials produced by AI, while the control group carried on with traditional teaching.

The project aimed at teaching English through visual, language, and PR activities. The images created with the help of AI were produced with regard to the learning targets of the lessons and applied as a source of vocabulary teaching, communication practicing, reading and writing.

6.4 AI Tools and Visual Materials

The AI images were beneficial for producing different images connected with characters, space, objects, deeds, feelings, and other important stories present in the lessons of English.

The instrument used for generating images was ChatGPT. The teacher formed the prompts according to the linguistic targets that he dealt with and evaluated the images before using them in the classroom.

The images in the form of visual prompts were applied separately and, at times, in succession to develop visual stories. The students were given the opportunity to interpret the presented images and discuss them as well as transform the visual information into spoken and written English..

6.5 Intervention Schedule

The intervention lasted eight weeks, with English lessons conducted according to the college's regular timetable. The experimental instruction followed a structured sequence:

Stage	Classroom Activity
1	Introduction of the topic and learning objective
2	Presentation of AI-generated visual material
3	Observation and interpretation of the visual
4	Vocabulary and language development
5	Speaking/discussion activities
6	Reading or text-image connection
7	Writing/creative response
8	Reflection and critical evaluation of the AI-generated visual

Each lesson lasted approximately 60 minutes, and approximately 4 lessons per week incorporated the intervention. The control group studied comparable English content during the same period but without the AI-generated visual-art intervention.

6.6 Intervention

The experimental group used computer-generated visual materials that were intended to fulfil the objectives of learning the English language. The activities of their lessons included:

Vocabulary development – naming and describing objects, people, places, actions, or feelings in the pictures.

Speaking – describing pictures, foretelling events happening, and talking about the characters.

Reading activities - matching images and texts and interpreting the images.

Writing activities – writing descriptions, dialogues, stories, or endings based on computer-generated images.

Grammar – using pictures to practise tenses, prepositions, adjectives, and sentence structures.

Creative and multimodal – interpreting visual information and turning it into spoken and written English.

The teacher helped students interpret the images and evaluate them critically.

6.7 Research Tools

Five instruments were utilized in the data collection process.

6.7.1 English Language Achievement Test

For the purpose of measuring students' overall language proficiency, an English language achievement test was conducted before and after the study.

6.7.2 Tasks for LSRW Assessment

For the assessment of listening, speaking, reading, and writing, separate tasks were created which were based on the objectives of the study.

6.7.3 Multimodal Literacy Assessment Rubric

A rubric for multimodal literacy assessment was used to assess students' ability to interpret visual information and convey meaning through visual and linguistic modes. The assessment curriculum included assessment of students' ability to interpret visual elements, make connection between visual and linguistic information, make meaning out of it, and communicate their interpretation.

6.7.4 Student Questionnaire

A questionnaire was designed and administered to the experimental group, whereby the students' experience and perception of AI-generated visuals were examined.

6.7.5 Classroom Observation

Classroom observation was employed during the lesson in order to document students' involvement and interaction in AI activities.

6.8 Validity and Reliability of Instruments

Before conducting the real study, the instruments used were analyzed for their appropriateness in terms of content validity. Two experts in the field of English Language Teaching/Applied Linguistics were consulted in order to review the instruments. Their comments were taken into account in order to improve the assessment instruments, tasks, and multimodality rubric.

6.9 Data Collection Procedure

In the course of the current investigation, data were collected in three stages.

Stage One: Pre-intervention

The subjects went through the procedure of taking the tests of achievement in English, determining the proficiency of their reading, listening, speaking and writing abilities, as well as assessing their skills of multimodality.

Stage Two: Intervention

The group under experimentation was receiving instruction in English with the use of images created by AI for eight weeks, while the control group had traditional instruction.

Stage Three: Post-intervention

In the end of the intervention, both groups completed the post-assessment tests. The subjects who were in the experimental group were given the questionnaire and the data from classroom observations were analyzed in order to find out the patterns concerning students' participation and involvement.

6.10 Ethical Principles

It was ensured ethical guidelines are followed by obtaining consent for conducting college-based research from participants prior to collecting data.

Also, the process of participation was made voluntary and identities of students did not reveal their identities. Collected data were allowed to be utilized only for research purposes. Additionally, images generated by AI were carefully assessed by a teacher to assure they do not produce inappropriate representation before being applied in classes.

6.11 Data Analysis

For the interpretation of data statistical methods that include descriptive and inferential statistics were utilized. Measures of central tendency were used when calculating means and SD values.

Also, paired samples t-tests were used to make comparisons between the performance of students in pre- and post-tests in each group while independent-sample t-tests were applied to compare the results of experimental and control groups. Additionally, ANCOVA was conducted to check how teaching method affects the results of post-tests without taking into consideration pre-test scores.

Prior to the inferential analyses, it was determined whether the following statistical assumptions were satisfied, i.e., normality, homogeneity of variance, linearity and homogeneity of regression slopes for ANCOVA. The level of statistical significance used was set at .05.

The qualitative data obtained from the questionnaire and classroom observations was analyzed thematically for identifying trends in students' perceptions and participation in AI-generated art.

7. RESULTS

The collected data from 100 persons were analyzed to find if AI-powered visual art had an influence on the process of learning English language and a marginal literacy growth. The processes of analyzing the received data included descriptive statistics, paired sample t-tests, independent sample t-tests, and analysis of variance. The effect sizes were calculated using Cohen's d and partial eta squares to estimate the effect of the observed effect on each group of respondents.

7.1 Descriptive Statistics

The means and the standard deviations were found for both groups both for all pre-test data and for post-test data, as well as for language skills and new media literacy. The findings of the data analysis are presented in Table 1.

7.2 LSRW Performance

The performance of the two groups across the four language skills is presented below.

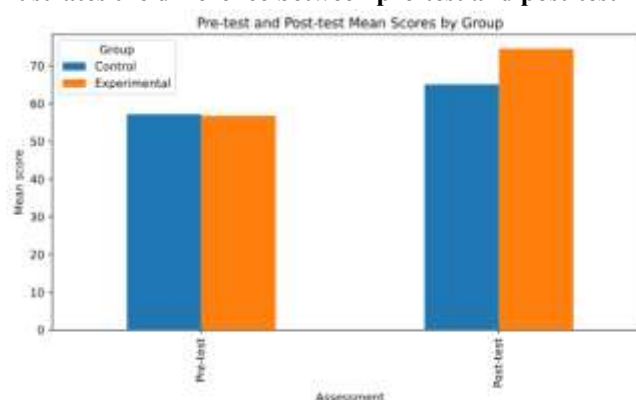
Group	<i>n</i>	Pre-test <i>M</i> (SD)	Post-test <i>M</i> (SD)	Listening <i>M</i> (SD)	Speaking <i>M</i> (SD)	Reading <i>M</i> (SD)	Writing <i>M</i> (SD)	Multimodal Literacy <i>M</i> (SD)
Experimental	50	56.80 (5.75)	74.60 (5.27)	75.20 (9.29)	73.80 (7.89)	76.09 (7.45)	73.21 (8.22)	76.40 (9.14)
Control	50	57.20 (6.25)	65.10 (5.78)	66.30 (8.29)	63.40 (8.07)	67.80 (8.48)	64.09 (7.78)	62.70 (10.29)

Note. *M* = mean; SD = standard deviation.

It can be seen from Table 1 that before the beginning of the program, the results of both the experimental and the control groups were fairly similar. The experimental group came out with an average of 56.80, while the control group had an average of 57.20. However, the post-test results show a significant divergence in favor of the experimental group, which had a score of 74.60 as opposed to the control group's score of 65.10. The same happens to listening, speaking, reading, writing, and multimodal literacy scores.

7.2 Comparison of Performance Before and After the Test

To examine changes in English language achievement within each group, paired-samples t-tests were conducted.

Figure 1 illustrates the difference between pre-test and post-test mean scores.

As illustrated in Figure 1, we can see that both groups of students showed progress after instruction but the gain made in the experimental group turned out to be much greater. The experimental group's average score improved from 56.80 to 74.60, while the average score in the control group increased from 57.20 to 65.10.

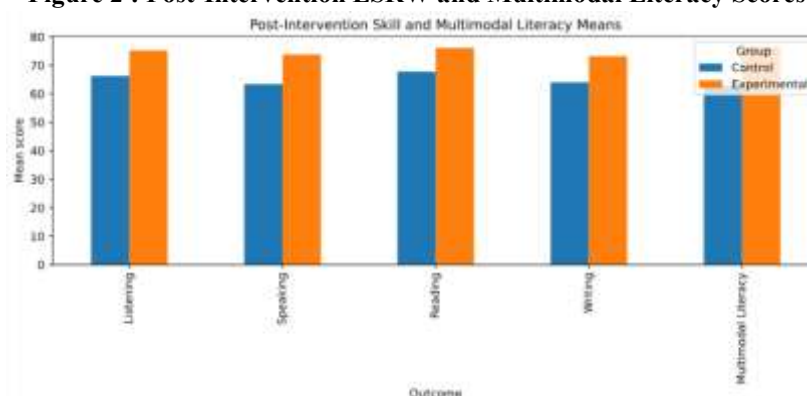
The paired-samples *t*-test results are presented in Table 2

Group	Pre-test <i>M</i>	Post-test <i>M</i>	Mean Difference	<i>t</i> (49)	<i>p</i>	Cohen's <i>d</i>
Experimental	56.8	74.6	17.8	19.46	< .001	2.75
Control	57.2	65.1	7.89	8.26	< .001	1.17

The experimental group exhibited notable progress as seen by the difference in pre-test and post-test, $t(49) = 19.46$, $p < .001$, $d = 2.75$. The control group similarly showed an improvement, $t(49) = 8.26$, $p < .001$, $d = 1.17$. The bigger effect size in the experimental group denotes a larger improvement from tests.

7.3 LSRW and Multimodal Literacy

The post-intervention performance of the two groups across the four language skills and multimodal literacy is presented in Figure 2.

Figure 2 : Post-Intervention LSRW and Multimodal Literacy Scores

According to the data illustrated in Figure 2, participants from the experimental group demonstrated higher average scores than their counterparts in the control group for all four domains of language proficiency tested. The area with the greatest discrepancy between the two groups is multimodal literacy, where the experimental group achieved an average of 76.40 compared to the control group's 62.70.

Independent samples *t*-tests were performed in order to establish the statistical significance of these distinctions. The relevant results are shown in Table 3.

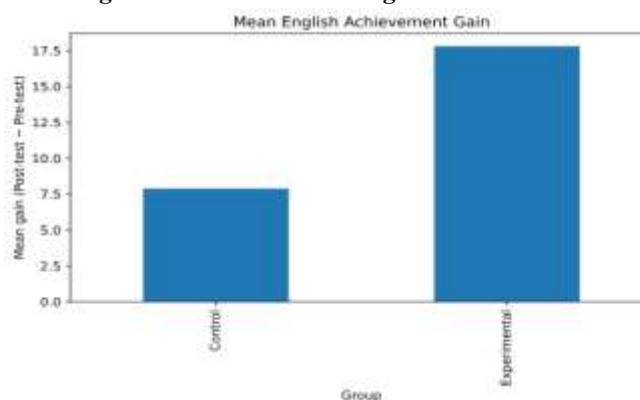
Table 3: Independent-Samples t-Test Results

Outcome	Experimental <i>M</i> (SD)	Control <i>M</i> (SD)	<i>t</i> (98)	<i>p</i>	Cohen's <i>d</i>
Post-test	74.60 (5.27)	65.10 (5.78)	8.6	< .001	1.72
Multimodal literacy	76.40 (9.14)	62.70 (10.29)	7.03	< .001	1.41
Post-test gain	17.80 (6.47)	7.89 (6.76)	7.49	< .001	1.5

The results showed that the experimental group had much better results on the overall test, $t(98) = 8.60$, $p < .001$, $d = 1.72$. There was also a statistically significant difference between the groups as far as multimodal literacy, $t(98) = 7.03$, $p < .001$, $d = 1.41$. This raises the conclusion that the differences observed in the study are not only statistically significant, but also indicate the magnitude of the effect.

7.4 Analysis of Gain Scores

To further examine the improvement resulting from the intervention, gain scores were calculated by subtracting each student's pre-test score from their post-test score.

Figure 3: Mean Gain in English Achievement

The experimental group demonstrated a mean gain of 17.80 points, whereas the control group demonstrated a mean gain of 7.89 points. The independent-samples t-test indicated that the difference between the groups was statistically significant, $t(98) = 7.49$, $p < .001$, with a large effect size ($d = 1.50$).

These findings indicate that students who participated in activities involving AI-generated visual art experienced substantially greater improvement in English achievement than students receiving conventional instruction.

7.5 Analysis of Covariance

Before the inferential analysis could be performed, it was checked whether certain statistical assumptions were fulfilled: normality, homogeneity of variance, linearity, and homogeneity of regression slopes with a level of significance of 0.05.

Qualitative data collected through the questionnaire and classroom observations have been systematically analyzed in order to detect the patterns in the students' attitudes and engagement in AI-generated art.

Table 4: ANCOVA of Post-test English Achievement by Instructional Group

Source	SS	df	MS	<i>F</i>	<i>p</i>	Partial η^2
Pre-test (Covariate)	356.446	1	356.446	13.095	< .001	0.119
Instructional Group	2,318.50	1	2,318.50	85.175	< .001	0.468
Error	2,640.39	97	27.221	—	—	—

Note. SS = sum of squares; MS = mean square; η^2_p = partial eta squared. Dependent variable = post-test English achievement; covariate = pre-test score.

The ANCOVA showed that the type of instruction has a significant influence on the students' English achievement at the post-test after controlling for their pre-test scores, $F(1, 97) = 85.175$, $p < 0.001$, partial $\eta^2 = 0.468$. The pre-test scores covariate was also statistically significant, $F(1, 97) = 13.095$, $p < 0.001$, partial $\eta^2 = 0.119$. The influence of instructional method is large, which means that we can see a difference in post-test results in the group of students who received AI-generated training compared to students who took traditional classes.

After adjusting for the pre-test scores, the mean results for the post-test were found to be 74.67 for the experimental group and 65.03 for the control group. Thus, following the results, it is possible to see that the experimental group still achieves significantly better results when we calculate the impact of their initial English proficiency level

7.8 Summary of the Findings

Overall, the findings imply that the experimental group performed consistently better on the primary measures in this study. The students exposed to AI generated visual arts performed better in English achievement as compared to their peers and hence statistically significant differences were seen in the post-test and gain scores. Improvement was also noted in the multimodal literacy levels of the experimental group.

While the paired sample t-test showed visible improvement in both groups, the experimental group showed the maximum impact. The independent sample t-test showed a notable variation between the groups in terms of achievement and multimodal literacy. Most importantly, ANCOVA confirmed that the difference in terms of post-test achievement remained significant after controlling for baseline scores.

Overall, it can be said that integrating AI-generated visual arts may serve as a meaningful multimodal strategy in teaching English.

8. DISCUSSION

The results of this research suggest a greater increase in the learners' performance of the English language while using the AI-generated art in comparison with traditional teaching. The experimental group decreased its mean score from 56.80 (SD = 5.75) during the pre-test to 74.60 (SD = 5.27) during the post-test while the control group ranged from 57.20 (SD = 6.25) to 65.10 (SD = 5.78). Using the independent-samples t-test proved there to be a statistically significant difference between groups in terms of post-test scores, i.e. ($t(98) = 8.60$, $p < .001$), which bears a significant magnitude of the effect ($d = 1.72$). ANCOVA showed the significance of the mode in teaching after controlling for the pre-test scores, $F(1, 97) = 85.18$, $p < .001$, partial $\eta^2 = .468$. Overall, the current data reflects the advancement of students taking part in the visual activities supported by AI as compared to students being traditional learners.

The enhancements in individual language skills corroborate this viewpoint. The means of language skills of the experimental group after the treatment were higher than the means of the control group in terms of listening (75.20), speaking (73.80), reading (76.09), and writing (73.21), while the means of the control group regarding the mentioned skills were 66.30, 63.40, 67.80, and 64.09 respectively. These statistics imply that it is possible to use AI-generated visual materials as a common field of experience, as described above. For example, students can hear descriptions of the picture while listening, then talk about what they see, followed by reading the relevant material, and finally writing a description or their version of the picture. Thus, the separation of language skills is eliminated, enabling learners to apply their knowledge of English for the purpose of communication.

The considerable progress made in multimodal literacy is noteworthy. The average score of the experimental group rose to 76.40 (SD = 9.14), whereas that of the control group was 62.70 (SD = 10.29). The results of the independent samples t-test showed a significant difference, $t(98) = 7.03$, $p < .001$, with a high effect size ($d = 1.41$). This indicates that the theoretical premise of the study suggesting the necessity of students' participation in meaning-making in different semiotic modes is confirmed. Kress (2010) states that human communication is now represented as a complex system of semiotic resources, including gestures, language, colour, images, and space. Furthermore, the visual arts created by AI allow the students to interact directly with various semiotic resources and make their interpretations through English.

The advancement of writing and speaking skills can also be attributed to the contextual and creative aspect of using visual prompts. A blank sheet can be a new challenge for students when it comes to coming up with new ideas, while an image can introduce specific characters, locations, actions, and events that will represent a good starting point for writing. As a result, in the current intervention, the participants had a chance to use visual sequences to create stories, describe characters, foresee events, and come up with different endings. Thus, visual materials were used as a tool for the beginning of the language production rather than its finale.

Results also point out to the impact on learners' engagement and creativity. Nearly all students from the experimental group had a good experience with the activities, with 88 of them confirming they found using images enjoyable and 90 of them believing that the activities led to greater creativity.

The results also underscore the importance of teacher intervention. Generative AI must not be considered a standalone educational tool. As AI-generated images may have misleading, stereotypical, and culturally unacceptable elements, UNESCO (Miao & Holmes, 2023) points to the need for a human-centred and ethical approach to applying Generative AI in education. A teacher's role in the English classroom is to manage the prompts, assess the content produced by AI, offer interpretation of the AI output, and promote questioning among the students instead of accepting the representation offered by the AI without questioning it.

The research results also add to emerging discussions about AI literacy in language education. It is essential for students to realize that an image produced by means of AI does not offer an unbiased picture of the world but rather represents an answer given by the AI based on its previous experience in processing similar requests made by the user. By asking students why the AI has presented a character, culture, place, or situation in a particular way, one can use a simple visual exercise to develop critical thinking skills among the students.

The outcomes also highlight the significance of guidance. It highlights that Generative AI should not be treated as an exclusive learning resource. Since AI imagery might be partial, stereotypical, and unacceptable from a cultural standpoint, UNESCO (Miao & Holmes, 2023) states that it is necessary to apply a human-oriented and ethical approach in the use of Generative AI in the process of education. The teacher's responsibility in an English class consists in providing instruction on how to prompt machines, evaluate the AI-generated material, explain it as well as encourage the pupils to question its credibility, rather than blindly accepting the provided solution. The findings can act as a contribution to discussions concerning AI literacy within language teaching. It is of great importance for learners to understand that a picture generated via AI does not reflect a true representation of reality, and it demonstrates the AI's response based on previous experience in processing the user's requests. By posing a question to students about the reasons for representing the characters/cultures/locations/situations in a particular manner, it is possible to implement a simple visual exercise in order to stimulate critical thinking among learners..

REFERENCES

1. Deng, L., & Jamaludin, K. A. (2026). Roles of Generative Artificial Intelligence (GenAI) in English as a Foreign Language (EFL) instruction: A systematic literature review. *SAGE Open*, 16(1). <https://doi.org/10.1177/21582440261418315>
2. Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
3. Li, B., Tan, Y. L., Wang, C., & Lowell, V. (2025). Two years of innovation: A systematic review of empirical generative AI research in language learning and teaching. *Computers and Education: Artificial Intelligence*, 9, 100445. <https://doi.org/10.1016/j.caeai.2025.100445>
4. Lee, S., Choe, H., Zou, D., & Jeon, J. (2025). Generative AI (GenAI) in the language classroom: A systematic review. *Interactive Learning Environments*, 335–359. <https://doi.org/10.1080/10494820.2025.2498537>
5. Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.
6. New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66(1), 60–93. <https://doi.org/10.17763/haer.66.1.17370n67v22j160u>
7. Yi, Y., & Ang, S. (2022). Multimodality in the English language classroom: A systematic review of literature. *Linguistics and Education*, 69, 101048. <https://doi.org/10.1016/j.linged.2022.101048>