ENHANCING ENGLISH LANGUAGE PROFICIENCY THROUGH MULTIMODAL LEARNING APPROACH: A SYSTEMATIC LITERATURE REVIEW

Grand Stef Yan Claudio[*1], Nadia Simare Mare[*2], Muhammad Subhan Fikri[*3], Tampung N. Saman[*4]

grandclaudio201@gmail.com[*1], nadiasimaremare0@gmail.com[*2], muhammadsubhanfikri@fkip.upr.ac.id[*3], tampung.saman@edu.upr.ac.id[*4]

Faculty of Teachers Training and Education[*1234]
University of Palangka Raya[*1234]

ABSTRACT

This study explores the application of multimodal approaches in English language learning and their effectiveness in improving English language skills. This systematic review analyzes 10 research articles published between 2021 and 2025 using the SALSA (Search, Appraisal, Synthesis, Analysis) method. A comprehensive search was conducted through electronic databases, including Google Scholar and ResearchGate, using specific keywords and predefined inclusion and exclusion criteria. The findings show that a multimodal approach positively contributes to language learning. These findings are categorized into 7 aspects of improvement, such as vocabulary, speaking, listening, reading comprehension and literacy, increasing student motivation and participation, supporting the integration of digital tools, and creating comprehensive multimodal strategies in the learning process. Overall, this systematic literature review concludes that multimodal learning approaches make a positive contribution that provides an effective, adaptive, and learner-centered strategy that enhances English language proficiency. However, most existing studies focus on short-term outcomes. Therefore, future research should examine long-term impacts, sustainability, diverse educational contexts, and the comparative effectiveness of specific multimodal strategies to provide deeper insight.

INTRODUCTION

Mastery of English as an international language has become an urgent and undeniable need in today's era of globalization. Isadaud et al. (2022) state that English, which is often used as a second language, is an international language that must be mastered by everyone because of its useful role in almost every aspect of people's lives. For this reason, English is a language that cannot be avoided in every aspect, such as technology development, education, or an international environment. Aini (2021) shows that the international language used

Grand Stef Yan Claudio[*1], *Nadia Simare Mare*[*2], *Muhammad Subhan Fikri*[*3], *Tampung N. Saman*[*4]

ENHANCING ENGLISH LANGUAGE PROFICIENCY THROUGH MULTIMODAL LEARNING APPROACH: A SYSTEMATIC LITERATURE REVIEW

to communicate in various countries is English. Not only as a means of communication, but English also functions as a bridge to access information, technology, and cultures from the outside world. However, in today's digital era, many students, especially in Palangka Raya, face difficulties in mastering English because they have not found the right learning method to improve their understanding of English.

The English learning method that is still widely used in many schools is still dominated by the conventional approach. This conventional approach includes a text-based system, which focuses on grammatical structure, translation, and memorization of vocabulary. This approach does not focus too much on the context of actual language use. Mahrlamova and Chabanovych (2021) argued that conventional learning is ineffective due to limitations in the application of media and technology and the lack of flexibility in time and place. This method also fails to meet the diverse and ever-changing learning needs of students. Furthermore, the learning system is one-way, preventing students from actively participating in the learning process. This can result in low student learning motivation, minimal involvement in the learning process, less than optimal learning outcomes, and a delayed language comprehension process.

In response to these challenges, a more inclusive, flexible, and relevant learning approach is needed. One learning approach that continues to develop and is starting to receive widespread attention is the multimodal learning approach. Kress (2010) states that multimodality refers to a concept that uses more than one mode of communication or various methods that can be used in a text to create meaning. Philippe et al. (2020) explained that the multimodal learning approach is a learning strategy that involves various forms of information representation, such as visual, auditory, kinesthetic, and textual, in the learning process. The multimodal learning approach refers to the use of various modes or channels in the learning process, such as audio, video, text, images, animation, body movement, and the use of digital technology development to facilitate and convey the material in the learning process. In other words, learning is not only delivered verbally, but also through visual, auditory, and kinesthetic representations simultaneously. Eisenmann and Summer (2020) added that the multimodal learning approach provides an opportunity for teachers to effectively pay attention to various learning styles at once. Multimodality in learning allows students to access information in more diverse and interesting ways, thus increasing their motivation and understanding.

English Language Proficiency (ELP) is an individual's skill in using English appropriately and efficiently to communicate in various situations, such as social, professional, and academic situations. Chapelle et al. (1997) stated that Language proficiency encompasses linguistic, sociolinguistic, and discourse knowledge as well as the ability to comprehend, employ, and also generate language in a range of communication circumstances. ELP covers several main language skills, namely listening, reading, speaking, and writing. ELP also provides a better understanding of vocabulary, language structure, and nuances of meaning in various social, professional, and academic situations. Sidabutar et al.

(2021) stated that language proficiency is often considered a measure of how fluent a person can use a language in various contexts.

In the context of this study, several variables are discussed include the multimodal learning approach, English language proficiency, and the impact received from using the multimedia learning approach. The multimodal learning approach is defined as the use of various media and methods in a learning process. Spring and Takeda (2024) stated that English language proficiency can be measured through listening, reading, speaking, and writing skills. The combination of these four skills will help students to understand the context of language learning. In the meantime, the influence is evaluated using the results of earlier research on this topic.

Many previous studies have discussed a multimodal learning approach. Some studies that are relevant to this topic include Firmansyah (2021), who stated that a multimodal approach can enhance students' multimodal competencies, such as data literacy, technology, and humanities, as well as strengthen understanding and analytical skills through various modalities such as language, images, music, sound, and movement. It also improves students' psychological analysis skills in the context of literature and language acquisition. A different study, by Sadikin et al. (2025), showed that pre-service teachers who were involved in digital multimodal composition practices experienced significant development in their narrative, grammar, and digital literacy skills.

Furthermore, continuous development of technology and the presence of artificial intelligence and interactive media have created a new space for language learning exploration that is no longer limited by the classroom. Papageorgiou and Lameras (2017) stated that the development of multimodality with the use of technology shows that technology supports various ways in a more active and collaborative learning process. In addition, technology also expands the possibilities for collaboration and exploration through multimedia and digital platforms. Fang (2025) further developed research on IMSEP (Intelligent Multimodal Spoken English Platforms) by designing a speech learning system based on the integration of CNN (Convolutional Neural Networks), DNN (Deep Neural Networks), and LSTM (Long Short-Term Memory). This platform enables multimodal English training equipped with adaptive feedback based on user performance. This proves that technology is not only a learning aid but can also function as a sophisticated pedagogical partner. On a global scale, the need for inclusive and responsive learning strategies to address the challenges of the times is increasingly pressing. Bekturova et al. (2025) conducted a study of EMI (English Medium Teaching) teachers in Kazakhstan and found that teachers' ability to utilize multimodal resources greatly determines the effectiveness of their teaching. Multimodality not only improves content understanding but also forms a more active and collaborative interaction dynamic in the classroom. Therefore, this approach also provides room for personalization of learning. Costa et al. (2025) developed a project in Portugal called the JASM project under the auspices of ESEV (Escola Superior de Educação de Viseu). This project focuses on foreign language learning through an approach that combines visual creativity and digital storytelling. The results of this project showed that multimodal media such as

Grand Stef Yan Claudio[*1], *Nadia Simare Mare*[*2], *Muhammad Subhan Fikri*[*3],
Tampung N. Saman[*4]
ENHANCING ENGLISH LANGUAGE PROFICIENCY THROUGH MULTIMODAL
LEARNING APPROACH: A SYSTEMATIC LITERATURE REVIEW

films, social media, and e-books were able to increase motivation, autonomy, as well as linguistic, cultural, and technological skills in foreign language learning. Maulidiya et al. (2023) found that interactive video-based multimodal learning media developed using H5P, guided by the ADDIE (Analysis, Design, Development, Implementation, and Evaluation) model was proven valid, interesting, and effective in improving students' speaking skills. Based on the validation results, the experts gave an assessment with a score above 92%, which indicates that the interactive video has a very high level of validity. This media has shown a clear impact in improvement the speaking ability skill of 10th-grade students with an index value of 0.548, which is classified as moderate. In addition, the level of student satisfaction with this media is quite high, reaching 81.25%, indicating that students are interested and enjoy the learning process using this media. However, to achieve more optimal results, further development is needed, which includes improving content, interactivity, and teacher support. The use of this media shows great potential in improving English learning and can be further developed for other subjects and other skills in the future. Rahmawanti et.al. (2021), through their research, showed that watching English films with subtitles is an effective and useful strategy in improving students' oral grammar skills. Both groups tested agreed that the application of multimodal video and text simultaneously or subtitled films helped them imitate language, expand vocabulary, improve pronunciation, and construct correct sentences. They also believe that watching films with subtitles repeatedly and voluntarily has a positive effect on understanding oral grammar, although the effect is not large. Overall, the use of subtitled English films can be an effective media and can be integrated into language learning methods to improve students' speaking competence. Meanwhile, Rohi and Nurhayati (2024) showed that multimodal learning strategies strengthen long-term memory, improve linguistic skills, creativity, and critical thinking skills.

An example of the implementation of a multimodal approach can be found in the use of multimodal texts. Lim et al. (2022) added that multimodal text consists of various semiotic resources, including words, images, symbols, and animations. These texts are often in the form of multimedia and digital. The examples of multimodal texts include posters, digital videos, collages, television advertisements, websites that provide various information, and technological communication. In learning, the multimodal texts commonly used are graphic novels or comic books. Lim et al. (2022) explained that graphic novels and comic books are created with integrated written text and images, which make easier for students to understand the meaning. Jasmine et al. (2025) added that the use of multimodal texts is approved in the improvement of literacy skills in English language learning. The research results showed that students have a high interest in learning about climate change issues through interactive visual, audio, and video media. This approach made learning more engaging and suitable, while also raising students' awareness of climate change and environmental issues.

With the advancement of technology and the need for more interactive learning, multimodal learning approaches are increasingly being researched and applied in English language teaching. However, previous studies on multimodal

approaches have largely focused on specific skills, including listening, writing, speaking, or reading, as well as on the use of specific technological tools or media such as audio, images, video, applications, or digital platforms. This fragmented approach has limited and fragmented research results, thus preventing comprehensive conclusions about how multimodal learning can effectively improve English language proficiency. Furthermore, very few studies have reviewed and synthesized previous findings through a systematic and structured synthesis process.

To address this gap, this study presents a Systematic Literature Review (SLR) with SALSA (Search, Appraisal, Synthesis, and Analysis) method to synthesize and organize findings from existing research on the implementation of multimodal learning approaches to improve English language proficiency. Through a comprehensive analysis, this study is expected to identify patterns and findings that offer new insight for educators and researchers in the future of the English language education environment.

The novelty of this study is the use of a comprehensive and holistic approach in reviewing the development of English language skills through multimodal learning. This study takes a holistic approach that encompasses all aspects of language skills, unlike previous studies that focused solely on one language skill or specific media. Furthermore, this study utilizes a rarely used method, the SALSA methodology, to ensure a systematic and transparent literature search, assessment, and analysis. Through this process, this study produces a synthesis of pedagogical strategies that integrate diverse media forms like audio, text, images, and video to design more interactive, effective, and engaging English language learning.

This study aims to explore and understand how a multimodal approach is applied in English language learning and to what extent it can improve students' language skills. Furthermore, this research also focuses on providing recommendations for educators to create and design curricula and learning strategies that are appropriate to students' needs. Through the results obtained, this study is supposed to offer academic and practical contributions for students, teachers, and the overall development of the quality of English language teaching and learning.

RESEARCH METHOD

The research design used in this study is the Systematic Literature Review (SLR). Triandini et al. (2019) stated that the SLR design is used to identify, review, evaluate, and interpret all available research on a topic area of interest, focusing on specific relevant research questions. The use of SLR design aims to identify the use of a multimodal approach in improving English language proficiency. This study used SALSA (Search, Appraisal, Synthesis, Analysis) method as a guide, originally developed by Booth et al. (2016). The SLR (Systematic Literature Review) design with SALSA (Search, Appraise, Synthesize, and Analyze) method is a structured way for researchers to collect, compile, and analyze literature systematically, comprehensively, and in a structured manner to explore and know the effectiveness of the multimodal learning approach in the improvement of

Grand Stef Yan Claudio[*1], *Nadia Simare Mare*[*2], *Muhammad Subhan Fikri*[*3], *Tampung N. Saman*[*4]

ENHANCING ENGLISH LANGUAGE PROFICIENCY THROUGH MULTIMODAL LEARNING APPROACH: A SYSTEMATIC LITERATURE REVIEW

English language competence. This SALSA method was chosen because it can capture the scope of the literature as a whole while providing depth of analysis in reviewing various approaches and research results that are relevant to the topic of improving English language skills through multimodal learning.

Search refers to the process of finding and identifying sources of information relevant to the topic or problem being researched. Search is the first stage that involves finding articles through various reputable academic databases. This stage is very important because it must ensure that all pertinent data and literature are found systematically and completely (Booth, 2016). The articles to be analyzed are identified through searches conducted on electronic databases, including Google Scholar and ResearchGate. Based on the search access, which found 63 articles about the use of multimodal learning.

The second stage of the SALSA framework is appraisal. Appraisal is a stage carried out to find information or data from previously published research articles. Appraisal is the process of assessing the quality, relevance, and reliability of sources of information or data that have been found during the search stage (Booth, 2016). The purpose of this stage is to ensure that only sources that meet the validity criteria and are included or excluded according to the specified criteria will be used in the next stage. The inclusion uses a combination of certain keywords. The keywords include multimodal learning approach, English language proficiency, the effect of using multimodal in speaking, listening, reading, writing, and multimodal technology for the learning process. Apart from the combination of keywords, the publishing limit is also set on publications within 5 years, namely, the range of 2021 - 2025. The five-year publication limit was applied to ensure that the reviewed literature encompasses contemporary advances and prevailing perspectives on the topic. These time parameters help to maintain the relevance, accuracy, and veracity of the investigation by focusing on recent research findings. The articles found represent the experiences and perceptions of language teachers and learners. The articles found were also written in English, with the type of document being a journal article. Meanwhile, the exclusions applied are articles written in Indonesian, review articles, articles explaining theory, and articles released more than 5 years ago. Based on search access, 43 articles were found on the use of multimodal approaches in learning. Based on the inclusion and exclusion criteria, 10 articles were selected for analysis and discussion.

Synthesis is the third stage, where all information from various literature sources is rearranged in a structured manner, so that it becomes easy to understand for analysis. Mengist (2020) argues that the synthesis stage is used to extract and categorize data based on certain themes relevant to the research topic. Booth (2016) added that in the synthesis section, there is a process of combining, grouping, and comparing findings from various studies to build a coherent narrative or structure. This synthesis process involves grouping the analyzed articles by various themes, including title and author name, year, participants, methods, and key findings. By doing this stage, the data or information found from the article becomes more organized so that it is easy to identify. By doing

this synthesis stage, the information that has been collected previously can be a strong basis for analysis in the next stage.

The last stage is analysis, which is the process of in-depth interpretation of the synthesis results. At this stage, a comprehensive analysis was carried out on the research findings related to the theory and context of English language education, while also examining the strengths and limitations of each approach identified in the literature. In addition, analysis is also used to find research gaps that can be used as a basis for further research. Grant and Booth (2009) state that analysis in the SALSA framework is the intellectual process of interpreting the results of synthesis to generate new knowledge or identify areas for further exploration. This stage helps to form critical conclusions or strong arguments from the research produced, especially in relation to the impact and role of multimodal learning in developing English language skills.

FINDING AND DISCUSSION
Finding
Various articles were found during the search and screening process. Ten articles were selected based on the inclusion and exclusion criteria. The analysis of these articles reveals that the use of multimodal learning approaches has a positive impact on the development of English language proficiency and comprehension of English as a Foreign Language (EFL). The findings of the analyzed articles are available in the table below.

Table 1. The Data of Articles with Inclusions Criterion

No.	Titles and Authors	Participants	Method	Findings
1.	Effectiveness of using multimodal approaches in teaching and learning listening and speaking skills Masinde et al., 2023	First-graders at Western Kenyan primary schools	Descriptive research with Mixed-methods	The research results show that the implementation of a multimodal approach is very effective for students because this approach accommodates students with different learning styles. Data from questionnaires, observations, and interviews revealed that students were more active in class when teachers used a multimodal approach compared to when teachers implemented a single approach using only textbooks as media.
2.	The role of multimodal learning in enhancing language acquisition for bilingual learners	Bilingual students in elementary and secondary school	Mixed-methods	The findings showed that multimodal strategies not only improve vocabulary acquisition and comprehension but also significantly increase learner motivation and memorization, making these methods indispensable in bilingual

	Wilson, 2024			education.
3.	The use of a digital multimodal textbook to foster students' English proficiency Aisyah & Nuraeni, 2022	16 students from Maros Muslim University's English Education Department's second semester of the 2022–2023 academic year	Quantitative research with a pre-experimental method	The results of the study demonstrated that students' English language competency was extremely low on the pre-test. Meanwhile, students were given treatment using digital multimodal textbooks in the post-test, and they showed greater improvement. The pre-test average score was 47.81 with a standard deviation of 14.61, according to the statistics. With a standard deviation of 13.54, the post-test average score rose to 65.50. Thus, after receiving treatment using digital textbooks, students' language proficiency improved.
4.	Enhancing student vocabulary proficiency through the effective implementation of multimodal approaches Nadia, 2023	22 students from Thammislam School in Thailand's fourth-grade primary school. Only class IV-A is utilized in this study.	Qualitative research	The findings indicate that the use of multimodal techniques, such as PowerPoint presentations, matching games, bingo games, Google Translate, and visual aids, has positively contributed to students' language learning progress and created a pleasant learning experience. Overall, the results of the interviews and observations indicated that using a multimodal approach to English learning can boost learning efficacy by about 80%.
5.	Multimodal learning material in an English-speaking class in Kampung Baluwarti Yuliani, 2022	English speaking class of the Kampung Baluwarti tourism community	Qualitative research	Most of the students have difficulty in applying their English orally. They have never practiced speaking in English, especially in running their business, and cannot speak well, or their English competence is still low. However, when multimodal learning materials are applied in learning in the Kampung Baluwarti English class, the learning materials are effective in improving the atmosphere of teaching and learning activities. The results showed that most students stated that the learning materials had attracted their attention and increased their motivation in

				learning English, especially in speaking.
6.	Multimodal - Based learning improves students' reading ability: In perspective of systemic functional linguistics Purba et al., 2023	210 Students and 9 Teachers in seven classes.	Quantitative research	The research results showed that the implementation of analytical learning systems in the classroom has not yet reached ideal standards. However, it has been demonstrated that junior high school students from a variety of backgrounds can enhance their analytical literacy abilities through the use of multimodal learning techniques, thus providing a more effective alternative for developing critical thinking skills and analytical understanding. By utilizing media such as picture books, videos, and digital sources that are adapted to the geographical context, culture, and school facilities, this approach is able to increase motivation, speaking and reading skills, and create a more positive learning atmosphere. The teacher's ability to apply this strategy effectively is a major factor in its effectiveness. Therefore, improving teacher competence and capacity is a key factor in ensuring that multimodal learning delivers maximum results for student learning development.
7.	Multimodal approach in teaching listening in EFL classroom of Indonesia: Teachers' perspective and experience Nata & Suharyanto, 2024	SMA N 1 Gemolong students	Qualitative research	This research has led to the successful implementation of multimodal methods for teaching listening skills. To enhance listening learning, teachers usually use interactive platforms like Google Classroom, Zoom, and visual PowerPoint presentations. Students gave various positive responses to several multimedia elements, considering them very interesting and effective for learning.
8.	The impact of multimodal learning on vocabulary acquisition among	60 students of the Al-Ikhlas MTs in Takkalasi in Class VIII	Quantitative research with a quasi-experimental	Research results show that integrating multimedia into EFL classes, as one of the most commonly used methods, has proven to have potential in the

				improvement of students' English language skills. The use of multimedia is considered to strengthen vocabulary mastery while enriching the learning experience. This approach creates a dynamic and more stimulating learning environment; thus, students are encouraged to participate more actively in the learning process.
	eighth-grade students Mukrim & Imran, 2025			
9.	The efficacy of collaborative and multimodal learning strategies in enhancing English language proficiency among ESL/EFL learners: A Quantitative analysis Salamanti et al., 2023	120 students from three UK secondary schools who are between the ages of 14 and 18.	Quantitative research	The research findings indicate that ESL/EFL students in UK secondary schools can improve their learning skills and engagement by using learning strategies that combine collaborative and multimodal learning. Based on the results of the independent sample t-test, the experimental group obtained significantly higher scores than the control group. These results indicate that students' engagement and motivation in learning English are positively influenced by the use of multimodal and collaborative learning approaches.
10.	Enhancing Students' Reading Comprehension of Recount Text with Multimodal Digital Literacy Mariam et al., 2025	Thirty-ninth graders at the Islamic secondary school in Semarang Regency in the 2024–2025 academic year	Qualitative research.	Through the implementation of the genre-based approach, which included stages such as BkoF (Building Knowledge of the Field), MoT (Modeling of the Text), JcoT (Joint Construction of the Text), and IcoT (Independent Construction of the Text), the approach has proven effective in improving students' comprehension, engagement, and ability to produce recount texts. Strategies such as text structure analysis, contextual learning, and the use of digital media (visual, audio, storytelling) make learning more interesting and improve outcomes, as seen from the increase in post-test scores (from 69 to 84). However, this approach has weaknesses, such as the need for a long time, a tendency to focus on rules, and difficulties for students with learning disabilities.

				Therefore, long-term success requires teacher collaboration, ongoing evaluation, and a balanced and culturally relevant approach.

From the analysis of all reviewed articles, the research findings can be grouped into several main themes, namely vocabulary improvement, speaking, listening, reading comprehension or literacy, digital tools, increasing motivation and participation, and comprehensive multimodal strategies.

Table 2. Articles Research Findings

No.	Category	Study
1	Vocabulary Improvement	Aisyah & Nuraeni (2022), Mukrim & Imran (2025), Nadia (2023)
2	Speaking	Masinde et al. (2023), Salamanti et al. (2023), Yuliani (2022)
3	Listening	Masinde et al. (2023), Nata & Suharyanto (2024), Salamanti et al. (2023)
4	Reading comprehension or literacy	Mariam et al. (2025), Purba et al. (2023), Wilson (2024)
5	Digital tools	Aisyah & Nuraeni (2022), Mariam et al. (2025), Nadia (2023), Salamanti et al. (2023)
6	Motivation and participation	Masinde et al. (2023), Purba et al. (2023), Salamanti et al. (2023), Yuliani (2022)
7	Comprehensive multimodal strategy	Mariam et al. (2025), Masinde et al. (2023), Purba et al. (2023), Salamanti et al. (2023), Wilson (2024)

Discussion

Vocabulary Improvement

Several studies reveal that the multimodal approach significantly enhances students' vocabulary mastery. Mukrim and Imran (2025) reported an increase in students' scores from 63.13 to 89.43 after applying visual and auditory integration, which helped strengthen word associations. Similarly, Nadia (2023) proved that using media such as Google Translate, pictures, and games called bingo and mix-and-match effectively supported Thai elementary students in overcoming language barriers and expanding their vocabulary. Aisyah and Nuraeni (2022) also showed that the use of a multimodal digital book (*Book Creator*), combining text, images, teacher voice recordings, and videos, positively affected students' basic vocabulary acquisition.

These findings reinforce Paivio's (1990) *Dual Coding Theory*, which explains that cognition involves two interconnected systems, namely verbal and nonverbal (imagery), that can enhance learning when activated together. In language learning, when learners simultaneously hear a word and see a related

ENHANCING ENGLISH LANGUAGE PROFICIENCY THROUGH MULTIMODAL
LEARNING APPROACH: A SYSTEMATIC LITERATURE REVIEW

image, two memory pathways are formed, reinforcing comprehension and retention. This is in line with the *Cognitive Theory of Multimedia Learning* by Mayer (2005), which states that students find it easier to understand by using pictures and words than from words alone. Consequently, the studies by Mukrim and Imran (2025), Nadia (2023), and Aisyah and Nuraeni (2022) confirm that multimodal learning strengthens vocabulary acquisition through integrated visual and auditory experiences.

Speaking Improvement
In the speaking context, Yuliani (2022) found that the use of moving images, sounds, and text encouraged students to be more confident and motivated to speak English in in Kampung Baluwarti, Surakarta. Likewise, Masinde et al. (2023) in Kenya discovered that first graders taught through multimodal methods (gestures, videos, and sound) became more active and engaged in simple conversations. Similarly, Salamanti et al. (2023) demonstrated that interactive exercises like multimedia-based group discussions and video presentations greatly enhanced students' speaking skills using collaborative-multimodal approaches.

This finding aligns with Vygotsky's (1978) Sociocultural Theory (as cited in Lantolf, 2000), which highlights how social contact and tools such as pictures, gestures, and videos in fostering language development. It is also consistent with Mayer's (2005) notion that multimodal presentations enhance student engagement and comprehension. Jewitt's (2010) argument clarifies that the use of multimodality can enrich learning interactions by providing multiple channels for meaning-making. Thus, learning strategies that combine multimodal and collaborative approaches are effective in enhancing students' confidence, motivation, and oral communication skills, particularly for those at the beginning level.

Listening Improvement
Multimodal approaches also play a crucial role in improving listening skills. Nata and Suharyanto (2023) found that high school students exposed to diverse media—films, audio clips, and visual presentations—understood oral information more effectively. This supports Paivio's (1990) *Dual Coding Theory*, which explains that information processed through verbal and nonverbal (visual) channels simultaneously enhances comprehension and memory.

Masinde et al. (2023) also emphasized that early childhood children understand language structure and context more easily through storytelling accompanied by gestures and pictures, consistent with Van Leeuwen's (2005) theory of multimodality, which posits that communication involves multiple modes beyond text, such as images, sound, and movement. Furthermore, Salamanti et al. (2023) found that regular audiovisual exposure improved both comprehension accuracy and students' concentration. This finding aligns with Sweller's (1994) *Cognitive Load Theory*, which suggests that multimodality reduces mental load by distributing information across different sensory channels, allowing for more efficient processing. Overall, these findings demonstrate that

multimodal learning strengthens listening comprehension, which is fundamental to overall language proficiency.

Reading Comprehension and Literacy Improvement
In literacy learning, Purba et al. (2023) revealed that a multimodal approach based on *Systemic Functional Linguistics* (SFL) enables students to interpret explicit and implicit meanings in texts using images, diagrams, and sounds. Lim et al. (2018) supported this, showing that SFL-based multimodal instruction effectively structures literacy learning at the secondary level. Mariam et al. (2025) also found an increase in students' scores from 69 to 84 after applying a digital, genre-based multimodal approach that incorporated historical videos, infographics, and digital storytelling. These findings align with Varaporn and Sitthitikul (2019), who reported that multimodal methods significantly improve adolescents' critical reading skills compared to conventional techniques. Wilson (2024) and Baharani and Ghafournia (2015) further confirmed that multimodal texts enhance bilingual literacy by integrating visual and auditory media. Thus, multimodality proves flexible and adaptable across linguistic and cultural contexts, supporting students' reading comprehension and cross-language literacy.

Digital Tools Usage
The advancement of educational technology has amplified the role of digital tools in learning, particularly multimodal learning. Research conducted by Aisyah and Nuraeni (2022) on the use of Book Creator showed positive results. Book Creator, used as a learning medium, combines several multimodal elements such as text, audio, video, and images in a single digital platform. Through this integration of various media, students not only receive information visually and verbally but also can directly engage with the material. This result is consistent with Mishra and Koehler's (2006) Technological Pedagogical Content Knowledge (TPACK) framework, which highlights the significance of combining technology, pedagogy, and content to enable effective teaching.

Nadia (2023) also showed that combining *PowerPoint*, *Google Translate*, and digital educational games effectively addressed language barriers in Thai schools. It is consistent with Vygotsky's (1978) sociocultural learning theory (Lantolf, 2000), which highlights the role of tools and social interaction in language development. Similarly, Salamanti et al. (2023) found that digital platforms such as *Padlet* and the *Digital Language Motivation and Engagement Scale* enhanced student engagement and collaboration. It is supporting Dörnyei's (2009) *Motivational Self-System Theory*, which links motivation to meaningful social activity. Furthermore, Mariam et al. (2025) reported that *YouTube* and *digital storytelling* effectively introduced recount texts in contextual and interactive ways, aligning with Bruner's constructivist perspective (Metsämuuronen & Räsänen, 2018) that emphasizes meaning-making through authentic learning experiences.

Grand Stef Yan Claudio[*1], *Nadia Simare Mare*[*2], *Muhammad Subhan Fikri*[*3],
Tampung N. Saman[*4]

ENHANCING ENGLISH LANGUAGE PROFICIENCY THROUGH MULTIMODAL
LEARNING APPROACH: A SYSTEMATIC LITERATURE REVIEW

Student Motivation and Participation Improvement

The significant rise in motivation and participation highlights that multimodal learning fosters more engaging and meaningful experiences. Yuliani (2022) observed that students who were previously passive became enthusiastic after the implementation of visual and interactive media. Similarly, Laksmi et al. (2021) found that the consistent use of interactive visual media, like audio and video, increased motivation in English learning.

Salamanti et al. (2023) reported that the experimental group's motivation level was 30% higher than the level of the control group. This is supported by Eugenijus (2023) statement that multimodal tasks combining audio, visual, and kinesthetic elements enhance enthusiasm and perseverance in language learning. Masinde et al. (2023) confirmed that first graders taught with multimodal methods were more active and responsive than those taught conventionally. Wati et al. (2023) showed that educational games improved vocabulary mastery, while Fitriani et al. (2024) and Purba et al. (2023) added that multimodal reading materials made students feel more comfortable and motivated. Overall, multimodal learning environments promote engagement, comfort, and active participation among learners.

Comprehensive Multimodal Strategy

Several researchers have developed comprehensive multimodal strategies integrating collaboration, genre-based learning, and semiotics. Salamanti et al. (2023) revealed that combining multimodality and collaborative learning enhanced students' speaking, listening, reading, and engagement skills simultaneously. This aligns with Hu et al. (2019), who stated that "collaborative multimodal tasks develop both linguistic competence and engagement simultaneously." Visual narrative projects involving videos, images, and texts also increased students' confidence by engaging them emotionally in meaning-making processes.

Wilson (2024) designed systematic multimodal strategies for bilingual learners combining technology, text manipulation, and social interaction. It is consistent with Cope and Kalantzis' (2015) *Multiliteracies Theory*, which emphasizes linguistic, cultural, and modal diversity in globalized learning contexts. Likewise, Purba et al. (2023) combined multimodality with SFL, showing that it deepens understanding of meaning in language structures. It is similar to Van Leeuwen's (2005) view that linguistic forms are inseparable from visual representation and social context.

Mariam et al. (2025) further demonstrated that combining genre-based approaches with digital storytelling improved students' critical thinking in interpreting and producing recount texts. Nugraha et al. (2024) similarly argued that "students develop a better understanding of textual purpose and audience when multimodal tasks are embedded in genre pedagogy." Masinde et al. (2023) also integrated *Multiple Intelligences Theory* and *Social Semiotics* into multimodal materials, addressing learners' cognitive and sensory diversity. This reflects Jewitt's (2010) assertion that multimodality provides learners with

multiple pathways for understanding, representation, and expression according to their individual strengths and contexts.

CONCLUSION AND SUGGESTION

The multimodal learning approach has a significant influence on the improvement of English language proficiency, according to the analysis and systematic literature evaluation. This approach has proven efficient in various aspects of language skills, including vocabulary mastery, speaking skills, listening skills, and reading comprehension or literacy. In addition to improving language skills, the multimodal approach also encourages students' participation and motivation in the learning process. Students feel more involved and interested in the material presented through various media. Therefore, they show increased concentration and learning outcomes. This demonstrates that education is not only more efficient but also more inclusive and tailored to each student's individual needs. Thus, it concluded that the implementation of a comprehensive multimodal learning strategy is highly recommended in English language learning. Through strategies and combinations of various media such as video, text, images, audio, interactive digital devices, and other media, students can learn in a more enjoyable, varied, and appropriate way to their respective learning styles.

However, it is important to realize that the results of this review have several limitations. The limited period of the study, between 2021 and 2025, may not be enough to offer the extensive picture of the long-term impact of this multimodal approach. In addition, the methods used in the studies analyzed are quite varied, which can affect the consistency of the results. Nevertheless, there are many interesting opportunities for further research. Long-term research is also needed to see the impact on students' English language skills over a longer period of time. With these considerations, a multimodal learning approach can be an important part of the transformation of English education towards a direction that is more pertinent to the needs of the times. Not just utilizing technology, but truly designing a learning experience that is more alive, meaningful, and by the unique characteristics of each student.

REFERENCES

Aini, M. R. (2021). Posisi bahasa Inggris sebagai bahasa kedua dan pembuatan lembar kerja peserta didik bagi siswa Smp/Mts se Kota Blitar. *Jurnal Abdimas Bela Negara*, *2*(1), 66–79. https://doi.org/10.33005/jabn.v2i1.40

Aisyah, S., & Nuraeni, N. (2022). The use of a digital multimodal textbook to foster students' English proficiency. *Jo-ELT (Journal of English Language Teaching) Fakultas Pendidikan Bahasa & Seni Prodi Pendidikan Bahasa Inggris IKIP*, *9*(2), 160. https://doi.org/10.33394/jo-elt.v9i2.6234

Baharani, B., & Ghafournia, N. (2015). The impact of multimodal texts on reading achievement: A study of Iranian secondary school learners. *International Journal of Applied Linguistics and English Literature*, *4*(4), 161–170. https://doi.org/10.7575/aiac.ijalel.v.4n.4p.161

Bekturova, M., Gaipov, D., Dyankova, G. H., & Tulepova, S. (2025). Examining multimodal literacy skills among EMI teachers in Kazakhstan. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE)*, *13*(1), 217–222. https://doi.org/10.23947/2334-8496-2025-13-1-217-222

Booth, A., Sutton, A., & Papaioannou, D. (2016). Systematic approaches to a successful literature review (1st ed.). *SAGE Publications.*

Chapelle, C., Grabe, W., & Berns, M. (1997). TOEFL: Communicative language proficiency: Definition and implications for TOEFL 2000. *Educational Testing Service.* http://www.toefl.org

Cope, B., & Kalantzis, M. (2015). The things you do to know: An introduction to the pedagogy of multiliteracies. In *a pedagogy of multiliteracies: Learning by design* (pp. 1–36). London: Palgrave Macmillan UK.

Costa, A. M., Costa, C. A., Coutinho, E., Oliveira, I., Pereira, J., Gillain, R., Amante, S., Fidalgo, S., Relvas, S., & Delplancq, V. (2025). Multimodal creativity at the service of foreign language teaching in higher Education in Portugal: The JASM project. *International Journal of Professional Business Review*, *10*(4), 1–11. https://doi.org/10.26668/businessreview/2025.v10i4.5361

Dörnyei, Z. (2009). The L2 motivational self system. *Motivation, language identity and the L2 self*, *36*(3), 9-11.

Eisenmann, M., & Summer, T. (2020). Multimodal literature in ELT: Theory and practice. *Children's Literature in English Language Education*, *8*(1), 52–73. https://storybird.com/

Eugenijus, L. (2023). The effectiveness of multimodal task design in second language teaching. *Research and Advances in Education*, *2*(12), 1–6. https://doi.org/10.56397/rae.2023.12.01

Fang, Q. (2025). Design of a future-oriented intelligent multi-modal spoken English platform: A study on deep integration of CNN-DNN-LSTM and self-supervised mechanisms. *Research Square*, 1–25. https://doi.org/10.21203/rs.3.rs-6366200/v1

Firmansyah, M. B. (2021). The effectiveness of multimodal approaches in learning. *Journal of Education and Technology*, *4*(3), 469–479.

Fitriani, S. A., Kurniawati, N., & Amuddin, M. (2024). Students' responses to multimodal texts in a learning reading course. *English Education: Jurnal Tadris Bahasa Inggris*, *17*(1), 148–164. https://doi.org/10.24042/ee-jtbi.v17i1.19459

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, *26*, 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Hu, W., Cai, B., Zhang, A., Calhoun, V., & Wang, Y. P. (2019). Deep collaborative learning with application to multimodal brain development study. *IEEE Transactions on Biomedical Engineering*, *66*(12), 3346–3359. https://doi.org/10.1109/TBME.2019.2904301

Isadaud, D., Fikri, M. D., & Bukhari, M. I. (2022). The urgency of English in the curriculum in Indonesia is to prepare human resources for global competitiveness. *DIAJAR: Jurnal Pendidikan Dan Pembelajaran*, *1*(1), 51–58. https://doi.org/10.54259/diajar.v1i1.177

Jasmine, M. P., Silvhiany, S., & Inderawati, R. (2025). Exploring young learners' needs in multimodal text design for integrating sustainable education into ELT. *Journal of Languages and Language Teaching*, *13*(2), 785–797. https://doi.org/10.33394/jollt.v13i2.13775

Jewitt, C. (2010). Technology and learning: A multimodal approach. In *International Handbook of Psychology in Education* (pp. 361–398).

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication* (13th ed.). Routledge. https://doi.org/10.4324/9780203970034

Laksmi, N. K. P., Yasa, I. K. A., & Mirayani, K. A. M. (2021). The use of animation video as learning media for young learners to improve EFL students' motivation in learning English. *Lingua: Jurnal Pendidikan Bahasa*, *17*(1), 42–52. https://doi.org/https://doi.org/10.34005/lingua.v17i1.1378

Lantolf, J. P. (2000). Introducing sociocultural theory. In *Sociocultural theory and second language learning* (pp. 1–26). Oxford University Press. www.oup.com/elt

Lim, F. V. (2018). Developing a systemic functional approach to teach multimodal literacy. *Functional Linguistics*, *5*(1), 1–17. https://doi.org/10.1186/s40554-018-0066-8

Lim, F. V., Toh, W., & Nguyen, T. T. H. (2022). Multimodality in the English language classroom: A systematic review of literature. *Linguistics and Education*, *69*. https://doi.org/10.1016/j.linged.2022.101048

Mahrlamova, K., & Chabanovych, N. (2021). Implementation of interactive methodology in medical education. *Linguistics and Culture Review*, *5*(S4), 1154–1160. https://doi.org/10.21744/lingcure.v5ns4.1752

Mariam, S., Kepirianto, C., Fadlilah, S., & Izza, A. F. (2025). Enhancing students' reading comprehension of recount text with multimodal digital literacy. *Indonesian EFL Journal (IEFLJ)*, *11*(1), 41–52. https://doi.org/10.25134/ieflj.v11i1.11337

Masinde, R., Barasa, D., & Mandillah, L. (2023). Effectiveness of using multimodal approaches in teaching and learning listening and speaking skills. *Nairobi Journal of Humanities and Social Sciences*, *7*(2), 91–112. https://doi.org/10.58256/njhs.v7i2.1333

Maulidiya, N. I., Makrifah, I. A., & Mubarok, T. A. (2023). Development of video-based interactive learning media to teach speaking skills at x grade students of SMK NU Donomulyo. *Wiralodra English Journal*, *7*(2), 53–64. https://doi.org/10.31943/wej.v7i2.234

Mayer, R. E. (2005). Cognitive theory of multimedia learning. *The cambridge handbook of multimedia learning* (pp. 31-48)

Mengist, W., Soromessa, T., & Legese, G. (2020). Method for conducting systematic literature review and meta-analysis for environmental science research. *Elsevier: MethodsX*, *7*(100777), 1–11. https://doi.org/10.1016/j.scitotenv.2019.134581

Metsämuuronen, J., & Räsänen, P. (2018). Cognitive-linguistic and constructivist mnemonic triggers in teaching based on Jerome Bruner's thinking. *Frontiers in Psychology*, *9*(2543), 1–13. https://doi.org/10.3389/fpsyg.2018.02543

Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record*, *108*(6), 1017–1054.

Mukrim, A., & Imran, A. F. (2025). The impact of multimodal learning on vocabulary acquisition among eighth-grade students. *Jurnal Pendidikan Bahasa Inggris Indonesia*, *13*(1), 11–18. https://doi.org/10.23887/jpbi.v13i1.4029

Nadia, R. (2023). Enhancing student vocabulary proficiency through the effective implementation of multimodal approaches. *Invention: Journal Research and Education Studies*, *4*(2), 33–40. https://doi.org/10.51178/invention.v4i2.1505

Nata, D. F. M., & Suharyanto, S. (2024). Multimodal approach in teaching listening in EFL classroom of Indonesia: Teacher's perspective and experience. *UMS Library*, 1–11.

Nugraha, I., Emilia, E., & Gunawan, W. (2024). A genre-based approach in teaching multimodal texts. *English Review: Journal of English Education*, *12*(3), 1093–1106. https://doi.org/10.25134/erjee.v12i3.9568

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press.

Papageorgiou, V., & Lameras, P. (2017). Multimodal Teaching and Learning with the Use of Technology: Meanings, Practices and Discourses. *In 14th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2017)* (pp. 133-140).

Philippe, S., Souchet, A. D., Lameras, P., Petridis, P., Caporal, J., Coldeboeuf, G., & Duzan, H. (2020). Multimodal teaching, learning, and training in virtual reality: A review and case study. *Virtual Reality and Intelligent Hardware*, *2*(5), 421–442. https://doi.org/10.1016/j.vrih.2020.07.008

Purba, R., Herman, H., Fatmawati, E., Saputra, N., & Sabata, Y. N. (2023). Multimodal-based learning improves students' reading ability in perspective of systemic functional linguistics. *Journal of Education Technology*, *7*(1), 61–70. https://doi.org/10.23887/jet.v7i1.543

Rahmawanti, R., Larasati, T. M., & Rahayu, N. (2021). Students' perception on the use of English-subtitled movies toward learner's spoken grammar understanding. Wira*lodra English Journal (WEJ)*, *5*(2), 39–49. https://doi.org/10.31943/wej.v5i2.135

Rohi, M. P., & Nurhayati, L. (2024). Multimodal learning strategies in secondary EFL education: Insights from teachers. *VELES JOURNAL: Voices of*

English Language Education Society, *8*(2), 458–469. https://doi.org/10.29408/veles.v8i2.26546

Sadikin, I. S., Santosa, I., Fadli, M. R., Fatonah, K., & Ulum, M. B. (2025). Exploring pre-service teachers' experiences with digital multimodal composing in narrative storytelling. *ELTIN JOURNAL: Journal of English Language Teaching in Indonesia*, *13*(1), 253–268.

Salamanti, E., Park, D., Ali, N., & Brown, S. (2023). Efficacy of collaborative and multimodal learning strategies in enhancing English language proficiency among ESL/EFL learners: A quantitative analysis. *Research Studies in English Language Teaching and Learning*, *1*(2). https://doi.org/10.62583/rseltl.v1i2.11

Sidabutar, U., Sinaga, N. T., Sitorus, N., & Lestari, F. D. (2021). A multimodal analysis on a vocational English book and its effect on students' English proficiency. *Linguistics and Culture Review*, *5*(S3), 1651–1665. https://doi.org/10.21744/lingcure.v5ns3.1951

Spring, R., & Takeda, J. (2024). Teaching phrasal verbs and idiomatic expressions through multimodal flashcards. *STEM: Journal of English Teaching through Movies and Media*, *25*(2), 40–53. https://doi.org/10.16875/stem.2024.25.2.40

Triandini, E., Jayanatha, S., Indrawan, A., Putra, G. W., & Iswara, B. (2019). Metode systematic literature review untuk identifikasi platform dan metode pengembangan sistem informasi di Indonesia. *Indonesian Journal of Information Systems*, *1*(2), 63-77. https://doi.org/10.24002/ijis.v1i2.1916

van Leeuwen, T. (2005). Multimodality, genre, and design. In *discourse in action* (pp. 73–93). Routledge.

Varaporn, S., & Sitthitikul, P. (2019). Effects of multimodal tasks on students' critical reading ability and perceptions. *Reading in a Foreign Language*, *31*(1), 81–108. http://nflrc.hawaii.edu/rfl

Wati, R. I., S, S., & Sah, R. W. A. (2023). Siswa sekolah dasar menggunakan game "New Family 100" untuk mengembangkan vocabulary. Bagaimana kegiatan implementasinya? *Jurnal Penelitian Tindakan Kelas*, *1*(2), 124–131. https://doi.org/10.61650/jptk.v1i2.220

Wilson, M. (2024). The role of multimodal learning in enhancing language acquisition for bilingual learners. *Emerging Education Research*, *2*, 1–11. https://emergingpub.com/index.php/er

Yuliani, F. (2022). Multimodal learning material in an English-speaking class in Kampung Baluwarti. *JETLE (Journal of English Language Teaching and Learning)*, *3*(2), 73–79. https://doi.org/10.18860/jetle.v3i2.15878