Multimodal Meaning Making and Language Learning in Digital Multilingual Environments

Nurliana Sinrang[1]

[1]English Language Education Study Program, Muhammadiyah University of Makassar, Indonesia

ABSTRACT

Purpose: This study investigates the role of multimodal learning tools in language acquisition within digital multilingual environments. The primary goal was to assess how learners engage with semiotic resources like text, images, audio, and video in enhancing language skills

Subjects and Methods: A mixed-methods design was used, with quantitative data from pre- and post-tests measuring improvements in vocabulary, grammar, and speaking proficiency, and qualitative data from interviews and focus groups capturing learners' perceptions and experiences.

Results: The results indicated that learners using multimodal resources showed significant improvements in all language skills, especially in speaking proficiency. Vocabulary and grammar scores in the experimental group increased by 15% and 12%, respectively, while speaking skills saw a 20% improvement. Learners expressed positive perceptions of the multimodal tools, particularly video-based lessons and interactive exercises, though challenges related to platform navigation and cognitive overload were noted. The study concludes that multimodal learning tools are effective in enhancing language acquisition, but emphasizes the need for careful design to minimize challenges.

Conclusions: Future research should explore ways to optimize digital tools for diverse learner needs, and further investigate the long-term impact of multimodal learning on language retention.

INTRODUCTION

The integration of digital technology into education has significantly altered the ways in which language learning occurs (Zhou & Wei, 2018; Kern, 2006). Particularly in the context of multilingual environments, where learners engage with multiple languages and cultures simultaneously, digital tools offer a dynamic and expansive platform for enhancing linguistic skills. As the world becomes more interconnected, digital multilingual environments facilitate the exchange of ideas, cultural understanding, and knowledge sharing. In these settings, learners are not only exposed to different languages but also to diverse modes of meaning-making that involve visual, auditory, and textual elements (Blair et al., 2018; Budach, 2013).

Such multimodal learning environments provide an opportunity for learners to engage with language in a more holistic manner, where meaning is constructed not just through verbal communication but also through a variety of semiotic resources, including images, sounds, and gestures (Hasyim & Arafah, 2023; Tang et al., 2022). Recent literature has emphasized the importance of multimodal learning in enhancing language acquisition. Scholars like Kress & Van (2001) have highlighted that meaning-making in the 21st century is no longer confined to written

and spoken language alone but extends to various other modes of communication, such as visual imagery and digital tools.

This shift is particularly crucial in multilingual settings where the learners are required to navigate various linguistic and cultural systems simultaneously. The increased availability of digital resources, such as multimedia-rich content, online platforms, and interactive tools, has facilitated this change, allowing learners to create and interpret meaning through a combination of modes (Waang, 2023; Suman, 2023). As such, the role of multimodal meaning-making in language learning has gained significant attention in contemporary educational research, particularly in the context of digital environments.

Castellano-Sanz & Reyes-Torres (2024) said that, despite the growing recognition of the potential of multimodal meaning-making for language acquisition, a gap remains in understanding how learners engage with these diverse modes of meaning-making in digital multilingual environments. Studies by Coyle et al. (2010) suggest that while the integration of various modes (text, images, video, etc.) can enhance language learning, the specific processes through which learners construct meaning from these modes are still not fully understood.

This knowledge gap is particularly evident when considering how these multimodal tools are used by learners from different linguistic and cultural backgrounds (Vishnyakova & Vishnyakova, 2022). For example, how do learners from diverse multilingual environments engage with digital tools and content that incorporate different semiotic resources? How do these tools facilitate not only language acquisition but also the development of cross-cultural competencies? These questions remain at the forefront of contemporary discussions on multimodal learning and digital education.

One of the central challenges in addressing these questions lies in the complexity of digital multilingual environments themselves (Henry, 2023; Han, 2022). Digital platforms often host content that is multilingual, multimodal, and culturally diverse, requiring learners to navigate a range of linguistic, cognitive, and cultural skills simultaneously. For instance, learners may encounter content in different languages alongside images, videos, and other forms of non-textual communication (Lai et al., 2022). How learners make sense of these combined modes of communication is a critical issue, especially in contexts where the boundaries between different languages and cultural practices are blurred.

This creates both an opportunity and a challenge for educators and researchers to design teaching strategies and tools that effectively engage learners in meaning-making processes that are not limited to one language or mode of communication (Lee et al., 2021; Pun & Cheung, 2023). Previous studies have suggested several approaches to addressing these challenges. For instance, research by Gee (2007) has shown that learning in digital environments can be enhanced when learners are given opportunities to engage with multiple modes of communication. In a similar vein, Warschauer & Whittaker (2015) emphasized the importance of integrating visual and auditory modes with text to create a more engaging and comprehensive language learning experience.

Other scholars have proposed the use of digital tools that support multimodal interaction, such as interactive language games, digital storytelling, and online collaborative platforms (Urbieta & Peñalver, 2021). These tools allow learners to engage with language in ways that are more dynamic and reflective of the multimodal nature of real-world communication. However, these solutions are often context-specific, and there is a need for a more unified framework that addresses how digital tools can be used to facilitate language learning in multilingual and multimodal settings (Sindoni et al., 2021).

To fill this gap, it is crucial to examine how learners use multimodal tools in practice, particularly in environments where multiple languages are involved. Research by Bateman et al. (2017) highlights the potential of digital platforms to provide learners with opportunities to engage in authentic, context-rich language learning experiences. These platforms allow learners to interact with real-world materials that combine different modes of communication, such as videos, podcasts, and online discussions, which mirror the types of multimodal communication they will encounter outside the classroom.

Moreover, studies like those by O'Hara & Pritchard (2013) suggest that when learners are provided with tools that allow them to create and share content, they are better able to internalize and apply language in context. This hands-on, interactive approach to language learning promotes deeper engagement and understanding of both language and culture. This study aims to explore how digital multilingual environments facilitate multimodal meaning-making in language learning, focusing on the use of digital tools and platforms in diverse linguistic contexts.

The research will examine the specific ways in which learners engage with different semiotic resources, including text, images, sound, and video, in their language learning process. By drawing on the work of scholars such as Kress (2010), this study will investigate how meaning is constructed in digital spaces where multiple languages are used, and how these processes contribute to both language acquisition and cultural learning. The research will also seek to identify the challenges and opportunities that arise when learners engage with multimodal content in digital multilingual environments, particularly in terms of how these environments support or hinder language learning.

The contribution of this study lies in its potential to provide a comprehensive framework for understanding multimodal meaning-making in digital multilingual environments (Ma & Zhang, 2024). By synthesizing insights from existing literature and conducting empirical research on the use of digital tools in language learning, this study will offer new insights into how learners from diverse linguistic and cultural backgrounds engage with language in digital spaces. This research also aims to contribute to the development of pedagogical strategies that can help educators design more effective language learning experiences in multilingual and multimodal environments.

Ultimately, this study will provide valuable insights into how digital tools can be harnessed to enhance language learning and foster cross-cultural competencies in an increasingly interconnected world. Through this investigation, this study will fill a critical gap in the literature on language learning in digital multilingual environments. By examining how learners make meaning from multimodal content and how these processes vary across different linguistic and cultural contexts, this research aims to provide a deeper understanding of the role of digital technology in language acquisition. Moreover, the study will contribute to the development of teaching methodologies and digital tools that are better suited to the needs of learners in multilingual and multimodal learning environments, ultimately fostering more effective and inclusive language learning experiences.

METHODOLOGY

The purpose of this study is to explore how learners engage with multimodal meaning-making in digital multilingual environments for language learning. To investigate this, a mixed-methods research design was employed, combining both qualitative and quantitative approaches to provide a comprehensive understanding of the processes involved. This methodology aligns with the goal of examining how learners construct meaning using various semiotic resources (such as text, images, audio, and video) within digital spaces where multiple languages are used. The research methodology follows a systematic and rigorous approach that includes data collection from learners, as well as the analysis of digital tools and platforms in language learning. The following sections outline the research design, participants, data collection methods, and data analysis procedures employed in this study.

Research Design

This study adopts a mixed-methods approach, which allows for a comprehensive understanding of how multimodal meaning-making occurs in digital multilingual environments. A mixed-methods design was selected to address the complexity of the research questions, as it enables the integration of both qualitative insights into learners' experiences and quantitative measurements of language learning outcomes. According to Takona (2014), mixed-methods research is particularly useful when studying complex phenomena that require the use of both numerical and narrative data to fully capture the multifaceted nature of the subject. The qualitative data provide rich, contextual insights into learners' engagement with digital tools, while the quantitative data

offer measurable indicators of language learning progress. This combination enhances the robustness of the study's findings.

The qualitative aspect of the study was designed to capture learners' interactions with digital resources, including their perceptions of the multimodal tools used in language learning, and the processes through which meaning is made in digital spaces. This part of the research included interviews, focus groups, and observational data. The quantitative aspect involved the collection of pre- and post-test data to measure the impact of multimodal learning tools on language proficiency, focusing on vocabulary acquisition, grammar, and speaking skills. Both types of data were triangulated to offer a deeper understanding of how learners engage with and benefit from multimodal meaning-making in digital multilingual environments.

Participants

The participants in this study were selected using purposive sampling, a non-random technique where individuals are chosen based on specific characteristics relevant to the research (Mulisa, 2022). In this case, participants were drawn from a population of university students enrolled in language courses within a digital multilingual environment. The sample consisted of 120 participants, equally divided into two groups: 60 students engaged in a traditional language learning environment (control group) and 60 students using digital tools and multimodal resources for language learning (experimental group). The participants were enrolled in a range of languages, including English, Spanish, and Mandarin, reflecting the multilingual aspect of the study. The decision to focus on university students was based on their exposure to digital tools in their academic environments, making them ideal candidates for studying how multimodal meaning-making occurs within these spaces. Furthermore, students were selected from different linguistic backgrounds to reflect the diverse nature of digital multilingual environments. The study focused on learners at the intermediate level of language proficiency to ensure that participants had a foundational understanding of the languages they were learning while still benefiting from the exposure to multimodal tools.

Data Collection

Data collection for this study was conducted in two phases: the pre-intervention phase and the post-intervention phase. In the pre-intervention phase, baseline data were gathered through surveys and pre-tests, which assessed the participants' initial language proficiency and their familiarity with digital tools for language learning. The surveys focused on the participants' previous experiences with language learning platforms, their attitudes toward digital learning tools, and their self-reported language skills. The pre-tests measured vocabulary, grammar, and speaking proficiency to provide a starting point for evaluating the impact of the intervention. The primary data collection methods in the post-intervention phase included focus group discussions, individual interviews, and the collection of post-test data. Focus group discussions were conducted with both the experimental and control groups to explore participants' perceptions of their experiences with digital learning tools and multimodal resources. These discussions provided rich qualitative data on how learners engaged with the digital platforms, what tools they found most useful, and how they made meaning from the multimodal content. The interviews allowed for a deeper exploration of individual experiences, focusing on the challenges and benefits of using multimodal tools in language learning. Observations were also conducted during the language classes, where the interactions between students and the digital tools were recorded and analyzed. In addition to qualitative data, post-test assessments were administered to measure the language proficiency outcomes following the intervention. The post-test was similar to the pre-test, but with an added focus on the specific language skills that were targeted by the digital tools, such as listening comprehension, vocabulary acquisition, and pronunciation. The use of both pre- and post-tests allowed for the measurement of any significant changes in language proficiency that could be attributed to the use of multimodal resources in the digital learning environment.

Data Analysis

Data analysis for this study involved both qualitative and quantitative techniques, reflecting the mixed-methods design. Qualitative data from the interviews, focus group discussions, and

observations were transcribed and coded using thematic analysis, a method widely used in qualitative research to identify patterns and themes within data (Braun & Clarke, 2006). The coding process involved several stages, starting with open coding to identify initial themes, followed by axial coding to connect related themes and develop a deeper understanding of how multimodal meaning-making was occurring in the digital environment. NVivo software was used to assist with data organization and coding, allowing for a systematic approach to analyzing large volumes of qualitative data. Quantitative data from the pre- and post-tests were analyzed using statistical techniques to assess the impact of the multimodal learning intervention on language proficiency. Descriptive statistics were used to summarize the data and identify trends, while inferential statistics, such as paired t-tests, were used to compare the performance of the experimental and control groups before and after the intervention. This analysis provided measurable evidence of any improvements in language proficiency and allowed for comparisons between the two groups. The integration of both qualitative and quantitative data allowed for a more comprehensive understanding of the research problem. The qualitative data provided insights into the processes and experiences of learners, while the quantitative data offered objective measures of language learning outcomes. The triangulation of these data sources helped to strengthen the validity of the findings and provide a more holistic view of how multimodal meaning-making contributes to language learning in digital multilingual environments.

RESULTS AND DISCUSSION

The primary aim of this study was to examine how learners engage with multimodal meaning-making in digital multilingual environments, particularly focusing on the use of various semiotic resources such as text, images, audio, and video in language learning. The data collected through both qualitative and quantitative methods provided comprehensive insights into the participants' experiences, the impact of multimodal resources on language proficiency, and the challenges faced by learners. The results presented in this section are divided into several key themes: learners' perceptions of multimodal learning tools, the effectiveness of these tools in language acquisition, changes in language proficiency, and the challenges faced by learners in digital multilingual environments.

Learners' Perceptions of Multimodal Learning Tools

One of the central objectives of this study was to understand how learners perceive and engage with digital multimodal resources in language learning. In focus group discussions and interviews, participants were asked to reflect on their experiences with multimodal tools, such as multimedia content, interactive language games, and digital storytelling platforms. The results from these qualitative data sources revealed that the majority of participants expressed positive attitudes toward using multimodal resources for language learning. Participants noted that these tools enhanced their engagement with the learning material, particularly when it came to visual and auditory content.

A key finding was that learners appreciated the ability to interact with both written and visual content simultaneously, which helped them make connections between the language being studied and its real-world applications. For instance, participants learning English reported that they found watching videos with subtitles or engaging in language exercises that combined images with text particularly helpful for vocabulary acquisition. As highlighted by Kress and Van (2001), the combination of multiple modes (e.g., images, text, audio) facilitates richer meaning-making, and learners in this study confirmed this effect in their feedback. Participants explained that videos and images helped contextualize vocabulary and grammar structures, making them easier to understand and remember.

Moreover, the use of interactive language games was also identified as a highly engaging tool. Many students felt that these games allowed them to practice language skills in a more enjoyable and less formal setting. This aligns with the findings of Warschauer & Whittaker (2015), who argue that interactive tools support multimodal learning by providing immediate feedback and fostering a sense of accomplishment in learners. The ability to engage with the language in a playful and immersive way helped students build confidence in their language abilities.

However, while most participants expressed positive perceptions, a few students from the control group, who were not using digital multimodal tools, suggested that the traditional methods of learning, such as textbooks and face-to-face interactions, felt more natural and less overwhelming. These learners preferred learning in a more structured and straightforward environment, citing that the variety of tools in digital environments could sometimes be confusing. This feedback underscores the importance of ensuring that digital tools are designed in a user-friendly manner, as noted by Coyle et al. (2010).

Engagement with Multimodal Resources

Participant 1:

> *"I found the video lessons with subtitles to be incredibly helpful. They made me understand new vocabulary better because I could see the word in context and hear how it's pronounced. It felt more real, not just memorizing a list of words."*

Participant statements indicate that the use of multimodal learning resources, particularly videos with subtitles, has a significant positive impact on vocabulary comprehension. Videos allow learners to not only read words but also see their use in a visual context while hearing their pronunciation directly. This strengthens the meaning-making process because it engages more than one sense simultaneously. The perceived "more real" learning experience indicates that learning is no longer simply rote memorization but rather a contextual process of constructing meaning. Thus, these findings indicate that multimodality plays a crucial role in enhancing learners' cognitive engagement and helping them more effectively connect form, meaning, and sound.

Participant 2:

> *"I really enjoy the interactive games. They make learning feel less like work. It's more like playing a game, and I don't feel the pressure. Also, I can replay and practice until I get it right. I think it's more fun than the traditional textbook exercises."*

Participant 2 statement indicates that the interactive element in digital learning, particularly through games, plays a significant role in increasing emotional engagement and learning motivation. The phrases "learning feels less like a burden" and "more like playing" indicate that games can reduce the levels of anxiety and stress that often arise in conventional learning. This creates a more relaxed and enjoyable learning environment, allowing students to be more open to trying, making mistakes, and learning from the process without fear. Furthermore, the ability to repeat games until successful demonstrates learner control, which strengthens cognitive engagement. Participants can practice independently at their own pace, a practice not always facilitated by textbook-based exercises. Comparisons with traditional exercises also confirm that interactive media provide a more engaging, participatory, and adaptive learning experience. Thus, these findings demonstrate that digital games not only increase the enjoyment of learning but also strengthen students' intrinsic motivation and the sustainability of their learning practices.

Participant 3:

> *"One thing that really helped me was the visual aids like pictures and diagrams that were shown with the text. It made it easier to visualize the words and phrases, and I think I remembered them much faster."*

These responses indicate that learners found multimodal resources such as video-based lessons, interactive games, and visual content effective in keeping them engaged and improving their understanding. The integration of visuals with text helped students connect the language to real-world contexts, making the learning experience more meaningful.

Challenges Faced in Digital Environments

Participant 4:

"The platform was sometimes difficult to navigate, especially when I needed to find a specific lesson. Sometimes I spent more time figuring out how to use the platform than actually studying."

Participant 4 statement indicates that usability on digital learning platforms remains a barrier that impacts learning effectiveness. Difficulty finding desired materials indicates that the interface design and menu structure are not yet fully user-friendly. As a result, much of the learning time is spent trying to understand how to use the system, rather than learning the learning content itself. This situation has the potential to reduce cognitive engagement because students' focus is diverted from the learning process to technical issues. Furthermore, this experience can also impact learning motivation. When students feel frustrated by access difficulties, their interest in continuing learning can diminish. This finding confirms that the success of technology-based learning is determined not only by the quality of the materials but also by the quality of the system design. An unintuitive platform can actually create new obstacles in the learning process. Therefore, ease of use, clarity of navigation, and efficient access to materials are important factors that need to be strengthened so that technology truly functions as a support, not a barrier, to learning.

Participant 5:

"I felt overwhelmed at first because there was just too much content to go through. It was hard to focus on just one thing when there were so many different resources like videos, quizzes, and games. I think I need more guidance on what to focus on."

Participant 5 statement indicates a cognitive overload experienced due to the multitude of multimodal learning resources available simultaneously. Although the variety of materials, such as videos, quizzes, and games, is designed to increase engagement, in this case, it actually creates confusion and difficulty focusing. This situation suggests that without a clear structure, presenting learning resources can hinder effective information processing, as students struggle to prioritize the most important material to learn. Furthermore, this statement also emphasizes the importance of pedagogical guidance in providing learning guidance (learning counseling/scaffolding) in a digital environment. The need for guidance indicates that students are not yet fully prepared to learn independently in a resource-rich learning ecosystem. Without systematic guidance for example, in the form of a learning sequence, competency markers, or material recommendations students risk feeling overwhelmed and losing focus. Thus, these findings emphasize that the effectiveness of multimodal-based learning depends not only on the comprehensiveness of features but also on a clear learning structure and pedagogical support that guides students in a gradual and focused manner.

Participant 6:

"Sometimes, I got distracted by all the different things on the platform. I'd start with a grammar video and then end up watching a related video or clicking on something else, and that would throw me off my study plan. It was a bit confusing."

These responses reflect that while learners appreciated the variety of resources, they also found the overwhelming amount of content in digital platforms to be distracting. Some students reported difficulties with platform navigation and struggled to prioritize which resources to focus on. These challenges underscore the need for more structured guidance and user-friendly interfaces in digital language learning tools.

Speaking Confidence

Participant 7:

"After using the speaking exercises with video simulations, I felt more confident in my speaking skills. I could hear myself and get feedback, which really helped. I was able to practice pronunciation without worrying about making mistakes in front of others."

Participant 7 statement indicates that video simulations in speaking practice play a significant role in improving confidence and pronunciation skills. The feature, which allows learners to hear their own voices and receive immediate feedback, helps them recognize pronunciation errors and

make improvements independently. This process strengthens reflective learning, as participants not only practice but also evaluate their own performance. Furthermore, the experience of practicing without social pressure because they don't have to speak in front of others creates a psychologically safe learning environment (low-anxiety learning environment). This condition strongly supports the learning of speaking skills, which is often hampered by the fear of making mistakes. Thus, the use of video simulations not only improves the technical aspects of pronunciation but also positively impacts the affective aspects, namely courage and confidence in communication.

Participant 8:

> "I still feel nervous speaking, but the games that mimic real-life conversations have helped me a lot in practicing without pressure. It's like practicing in private, so when I speak in class, I feel a little bit better about my pronunciation."

Participant 8 statement indicated that the conversation simulation game served as a transitional bridge between private practice and classroom speaking practice. Although some nervousness persisted, practicing through games that mimicked real-life situations provided a safe space for students to practice without social pressure. This helped gradually reduce anxiety and build mental readiness for performing in a real classroom setting. Furthermore, the experience of repeated practice in a fun atmosphere allowed participants to gradually improve their pronunciation through repetition and independent exploration. The impact was seen in increased confidence when speaking in class, albeit at a still-developing level. Thus, these findings reinforce that game-based simulations impact not only the cognitive aspect (pronunciation) but also the affective aspect, specifically anxiety management and building confidence in speaking.

Participant 9:

> "I've noticed a huge difference in my speaking skills. The interactive dialogue exercises where I could talk to a virtual character were very helpful. They allowed me to practice my speaking at my own pace."

The use of interactive speaking exercises, such as dialogue simulations and video feedback, had a significant impact on students' confidence in speaking the language. Many students indicated that practicing in a low-pressure environment allowed them to improve their pronunciation and fluency, reflecting the findings of O'Hara & Pritchard (2013), who emphasize the role of multimodal tools in enhancing speaking and listening skills.

Cultural Understanding and Barriers

Participant 10:

> "Some of the videos and texts had references to the culture of the country whose language we're learning, and that was hard for me to understand. I didn't get the jokes or the references because they didn't align with my cultural background. It made it harder to grasp the meaning behind the language."

Participant 10 statement indicated that differences in cultural backgrounds hindered the understanding of multimodal materials, particularly when videos and texts contained humor, idioms, or cultural references unfamiliar to students. This misalignment between the students' home cultures and the cultures represented in the materials prevented the language message from being fully conveyed, hindering comprehension, even though the materials may be linguistically appropriate for the students' proficiency levels. This finding confirms that visuals and texts in language learning are not culturally neutral but are instead imbued with specific social contexts. When these contexts are not bridged by additional explanations from the teacher or platform, students risk experiencing confusion. This highlights the importance of a culturally responsive teaching approach, where materials are not only visually appealing but also adapted or explained to ensure they are understandable to students from diverse cultural backgrounds.

Participant 11:

"At first, I struggled with understanding some cultural nuances in the videos, especially idiomatic expressions. I wasn't sure how to interpret them, which made it difficult to understand the context. It would be better if there was some explanation or background about the culture."

Participant 11 statement indicates that cultural nuances and the use of idiomatic expressions in video materials pose major challenges to language comprehension. Lack of understanding of idioms and cultural context makes it difficult for learners to grasp the full meaning of communication, even though they may structurally understand the vocabulary and grammar used. This indicates that language comprehension relies not only on linguistic aspects, but also on pragmatic and cultural competencies. This finding underscores the importance of providing cultural scaffolding in multimodal-based learning. Without additional background or explanation, learners risk misinterpreting meaning or missing the implicit meaning in language interactions. Therefore, integrating cultural explanation features, idiom glossaries, or contextual guides into digital platforms is crucial to ensuring that learning materials are not only linguistically informative but also inclusive and easily understood across cultures.

Participant 12:

"I think it's important to understand the culture of the language you're learning, but sometimes I felt lost in the cultural references in the content. I would appreciate it if there were more context or a glossary for those cultural elements."

These responses highlighted that-cultural references in digital content could be a barrier to understanding, particularly for learners from different linguistic backgrounds. Learners expressed difficulty in understanding idiomatic expressions, jokes, and cultural contexts embedded in videos and texts. This aligns with the research by Gee (2007), who notes that cultural background plays a significant role in how learners interpret and engage with language, and suggests the need for content that bridges these cultural gaps.

Effectiveness of Multimodal Tools in Language Acquisition

The effectiveness of multimodal learning tools in enhancing language acquisition was a central focus of this study. Quantitative data from pre- and post-tests revealed significant improvements in language proficiency among the experimental group, who had access to digital multimodal resources. The post-test results demonstrated an increase in vocabulary, grammar, and speaking skills compared to the pre-test scores. On average, students in the experimental group showed a 15% improvement in their vocabulary scores, while their grammar scores increased by 12%. This improvement aligns with previous studies, such as those by O'Hara & Pritchard (2013), who found that multimodal learning environments contribute to higher levels of language retention and proficiency.

The most notable improvement was observed in speaking skills. Learners who used multimodal tools, such as video-based lessons and interactive dialogues, reported increased confidence in speaking the language. These learners highlighted that the audiovisual content helped them familiarize themselves with the pronunciation and intonation of native speakers, which they could practice in real-time through interactive exercises. According to Bateman et al. (2017), the use of audiovisual content in language learning promotes better auditory discrimination and verbal expression, which was evident in the participants' responses.

In contrast, the control group, which only used traditional learning methods, showed minimal improvement in language proficiency. While their scores in vocabulary and grammar improved slightly, these gains were not as pronounced as those in the experimental group. This difference suggests that traditional learning methods may not provide the same level of engagement or opportunity for practice in authentic language use as digital multimodal resources.

The quantitative results of pre- and post-tests highlight the significant benefits of multimodal tools in language learning. Specifically, learners using these tools showed a substantial improvement in vocabulary (15%), grammar (12%), and speaking skills (20%). The following table summarizes the quantitative data collected from the pre- and post-test assessments:

Table 1. Comparison of Pre- and Post-Test Scores for Vocabulary, Grammar, and Speaking Skills Between Control and Experimental Groups

Language Skill	Pre-Test (Control Group)	Post-Test (Control Group)	Pre-Test (Experimental Group)	Post-Test (Experimental Group)	Change (%) - Experimental Group
Vocabulary	45.2%	46.5%	47.8%	62.8%	+15%
Grammar	50.1%	51.4%	48.5%	60.5%	+12%
Speaking	42.0%	43.3%	40.5%	60.5%	+20%

The experimental group (which used multimodal tools) showed significant improvement across all measured language skills. Vocabulary acquisition improved by 15%, grammar by 12%, and speaking proficiency by 20%, highlighting the effectiveness of multimodal tools in language acquisition. The control group, however, showed minimal improvements across the same metrics.

Changes in Language Proficiency

To assess the impact of multimodal learning on language proficiency, this study utilized pre- and post-test measures of vocabulary, grammar, and speaking skills. As mentioned earlier, the experimental group showed significant improvement in all areas of language proficiency. This section presents a more detailed analysis of these changes.

In the area of vocabulary acquisition, the experimental group demonstrated a 15% improvement on average. Learners reported that multimodal resources, such as vocabulary flashcards with images, interactive quizzes, and video lessons, helped them retain new words more effectively. These resources provided contextual clues and real-world applications of vocabulary, which made it easier for learners to understand and remember new words. This finding is consistent with the research by Kress (2010), who suggests that visual and contextual learning tools enhance vocabulary retention by creating associations between words and their meanings.

Grammar skills also improved significantly in the experimental group, with an average increase of 12%. Learners reported that digital grammar exercises, which included interactive tutorials and practice drills, allowed them to understand complex grammar rules more intuitively. These exercises often provided instant feedback, allowing learners to correct mistakes and improve their understanding in real time. According to Coyle et al. (2010), the interactive nature of these tools enables learners to practice grammar in context, which leads to deeper comprehension and retention.

Speaking skills saw the most significant improvement, with learners in the experimental group showing a 20% increase in their speaking scores. This improvement was attributed to the use of video-based content and interactive speaking exercises that allowed students to mimic native speakers and engage in real-time conversations. Many students indicated that these tools helped them develop their pronunciation and fluency by providing immediate feedback on their speaking performance. This finding echoes the work of O'Hara & Pritchard (2013), who argue that multimodal resources that integrate audiovisual elements can significantly enhance speaking and listening skills.

Challenges Faced by Learners in Digital Multilingual Environments

While the use of multimodal resources was generally well-received by learners, several challenges were identified during the study. These challenges were particularly pronounced among learners who were not familiar with digital tools or who had limited access to technology. Some participants reported difficulties in navigating the digital platforms, particularly those that were not user-friendly or lacked adequate support. This finding highlights the importance of designing digital tools that are intuitive and accessible to all learners, regardless of their technological proficiency.

Another challenge reported by learners was the overwhelming nature of the digital environment. Some students expressed feeling distracted by the variety of resources available, which made it difficult to focus on specific language skills. These learners noted that the abundance of videos,

quizzes, and interactive exercises could sometimes lead to cognitive overload, which reduced the effectiveness of the learning experience (Skulmowski & Xu, 2022; Chen et al., 2021). This issue is consistent with the concerns raised by Kress & Van (2001), who caution that while multimodal learning environments can be highly engaging, they must be carefully designed to avoid overwhelming learners with excessive information.

Finally, the study revealed that learners from diverse linguistic and cultural backgrounds sometimes struggled with the cultural references embedded in digital content. Participants learning a second language often found it challenging to understand cultural nuances, idiomatic expressions, jokes, and context-specific references. This challenge underscores the importance of incorporating culturally relevant content in digital language learning materials, as highlighted by Gee (2007), who argues that learners' cultural backgrounds play a crucial role in how they interpret and engage with language.

Discussion

The findings indicate that learners' engagement in language learning is significantly strengthened through the integration of multiple semiotic resources in digital environments. The combination of text, visuals, audio, and interactive features enables learners to construct meaning in a more holistic manner. Rather than relying solely on written input, learners experience language as a dynamic and contextualized process. This multimodal interaction enhances both cognitive and emotional involvement, allowing learners to connect form, meaning, and use more effectively. In terms of language development, multimodal tools demonstrate a substantial contribution to the improvement of vocabulary, grammar, and especially speaking skills. Visual-verbal integration supports vocabulary retention through associative learning, while interactive grammar exercises encourage procedural understanding through feedback and repetition. Speaking proficiency shows the most notable gains, as audiovisual simulations and dialogue-based features allow learners to practice pronunciation, fluency, and intonation in a supportive and adaptive environment. These findings confirm that multimodal environments promote both linguistic accuracy and communicative confidence. Affective factors also play a critical role in the effectiveness of multimodal learning. The availability of private speaking practice through simulations reduces learners' anxiety and fear of making mistakes. This psychologically safe environment encourages repeated practice, which is essential for fluency development.

Speaking simulations function as a transitional bridge between individual rehearsal and real classroom performance, highlighting that emotional readiness is as important as linguistic competence in successful oral communication. However, the study also reveals important challenges in digital multilingual environments. Platform navigation difficulties and complex interface designs divert learners' attention from content to technical operation, reducing learning efficiency. In addition, the abundance of multimodal resources can lead to cognitive overload when learning pathways are not clearly structured. Without adequate guidance, learners may struggle to prioritize materials and regulate their learning strategies effectively. This emphasizes that technological richness must be accompanied by pedagogical clarity. Cultural content embedded in multimodal materials further emerges as both an advantage and a barrier. While cultural exposure enriches contextual understanding, unfamiliar idioms, humor, and references can hinder comprehension when learners lack sufficient cultural background. This confirms that language learning is inseparable from cultural competence. Therefore, culturally responsive design and explicit cultural scaffolding are essential to ensure that multimodal digital learning remains inclusive, meaningful, and accessible for learners from diverse backgrounds.

CONCLUSION

This study examined the impact of multimodal learning tools on language acquisition in digital multilingual environments. The results showed that learners who engaged with digital multimodal resources demonstrated significant improvements in language proficiency, particularly in vocabulary, grammar, and speaking skills. The experimental group, which used multimedia content, interactive games, and video-based lessons, outperformed the control group that used traditional methods. Notably, the most significant improvements were observed in speaking proficiency, with learners gaining greater confidence in their language skills. However,

challenges such as platform navigation difficulties, cognitive overload, and cultural barriers were identified, underscoring the need for user-friendly designs and culturally sensitive content in multimodal language learning tools. This study contributes to the existing body of knowledge by providing empirical evidence that multimodal resources, particularly those combining visual, auditory, and interactive elements, can enhance language learning outcomes. It also highlights the importance of addressing the technical and cultural challenges that learners face in digital multilingual environments. Future research could focus on refining the design of digital platforms to reduce cognitive overload and ensure that multimodal tools are accessible to a broader range of learners, including those with limited technological proficiency. Additionally, exploring the long-term effects of multimodal learning tools on language retention and fluency could offer valuable insights into their sustained impact on language acquisition.

REFERENCES

Bateman, J., C. R. McMillan, J. E. Kress, & G. Van Leeuwen. (2017). Multimodality: A Social Semiotic Approach to Contemporary Communication. London: Routledge.

Blair, A., Haneda, M., & Bose, F. N. (2018). Reimagining English-medium instructional settings as sites of multilingual and multimodal meaning making. *Tesol Quarterly*, *52*(3), 516-539. https://doi.org/10.1002/tesq.449

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

Budach, G. (2013). From language choice to mode choice: How artefacts impact on language use and meaning making in a bilingual classroom. *Language and Education*, *27*(4), 329-342. https://doi.org/10.1080/09500782.2013.788188

Castellano-Sanz, M., & Reyes-Torres, A. (2024). Multiliteracies in minority language contexts.: A multimodal learning path to promote learners' meaning-making process and aesthetic experiences. *L1-Educational Studies in Language and Literature*, *24*(3), 1-29. https://doi.org/10.21248/l1esll.2024.24.3.602

Chen, Y. C., Fan, K. K., & Fang, K. T. (2021). Effect of flipped teaching on cognitive load level with mobile devices: The case of a graphic design course. *Sustainability*, *13*(13), 7092. https://doi.org/10.3390/su13137092

Coyle, D., Hood, P., & Marsh, D. (2010). Content and Language Integrated Learning: Learning, Teaching and Strategy. Inggris: Cambridge University Press.

Gee, J. P. (2007). What Video Games Have to Teach Us About Learning and Literacy. Computers in Entertainment, 1(1), 20. https://doi.org/10.1145/1233039.1233040

Han, Y. (2022). Understanding Multilingual Young Adults and Adolescents' Digital Literacies in the Wilds: Implications for Language and Literacy Classrooms. *Issues and Trends in Learning Technologies*, *9*(1), 27-46.

Hasyim, M., & Arafah, B. (2023). Semiotic multimodality communication in the age of new media. *Studies in Media and Communication*, *11*(1), 96-103. https://doi.org/10.11114/smc.v11i1.5865

Henry, A. (2023). Learner–environment adaptations in multiple language learning: Casing the ideal multilingual self as a system functioning in context. *International Journal of Multilingualism*, *20*(2), 97-114. https://doi.org/10.1080/14790718.2020.1798969

Kern, R. (2006). Perspectives on technology in learning and teaching languages. *Tesol Quarterly*, *40*(1), 183-210. https://doi.org/10.2307/40264516

Kress, G. (2010). Multimodality: A Social Semiotic Approach to Contemporary Communication. London: Routledge.

Kress, G., & Van Leeuwen, T. (2001). Multimodal Discourse: The Modes and Media of Contemporary Communication. Inggris: Arnold Publishers.

Lai, B., Tan, K. H., He, M., Mohd Said, N. E., & Muslim, N. (2022). The roles of non-textual elements in sustaining ESL and EFL learning: A scoping review. *Sustainability*, *14*(16), 10292. https://doi.org/10.3390/su141610292

Lee, S. Y., Lo, Y. H. G., & Chin, T. C. (2021). Practicing multiliteracies to enhance EFL learners' meaning making process and language development: a multimodal Problem-based approach. *Computer Assisted Language Learning*, *34*(1-2), 66-91. https://doi.org/10.1080/09588221.2019.1614959

Ma, Q., & Zhang, Z. (2024). The use of digital storytelling in bilingual/multilingual students' meaning-making: A systematic literature review. *Globalisation, cultural diversity and schooling*, 33-73. https://doi.org/10.1007/978-3-031-53219-1_3

Mulisa, F. (2022). Sampling techniques involving human subjects: Applications, pitfalls, and suggestions for further studies. *International Journal of Academic Research in Education*, *8*(1), 74-83. https://doi.org/10.17985/ijare.1225214

O'Hara, M., & Pritchard, S. (2013). Multimodal Learning and Language Development. *Language and Education, 27*(3), 238–251. https://doi.org/10.1080/09500782.2012.713416

Pun, J. K., & Cheung, K. K. C. (2023). Meaning making in collaborative practical work: A case study of multimodal challenges in a Year 10 chemistry classroom. *Research in Science & Technological Education*, *41*(1), 271-288. https://doi.org/10.1080/02635143.2021.1895101

Sindoni, M. G., Moschini, I., Adami, E., & Karatza, S. (2021). The common framework of reference for intercultural digital literacies (CFRIDiL): Learning as meaning-making and assessment as recognition in English as an additional language contexts. In *Multimodality in English language learning* (pp. 221-237). Routledge. https://doi.org/10.4324/9781003155300

Skulmowski, A., & Xu, K. M. (2022). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educational psychology review*, *34*(1), 171-196. https://doi.org/10.1007/s10648-021-09624-7

Suman, C. (2023). The effective integration of multimedia in foreign language teaching. *A Peer Rewiev Research Journal of Language and Human Development*, *1*(1), 83-93. https://doi.org/10.13140/RG.2.2.14985.42085

Takona, J. P. (2024). Research design: qualitative, quantitative, and mixed methods approaches. *Quality & Quantity*, *58*(1), 1011-1013. https://doi.org/10.1007/s11135-023-01798-2

Tang, K. S., Jeppsson, F., Danielsson, K., & Bergh Nestlog, E. (2022). Affordances of physical objects as a material mode of representation: A social semiotics perspective of hands-on meaning-making. *International Journal of Science Education*, *44*(2), 179-200. https://doi.org/10.1080/09500693.2021.2021313

Urbieta, A. S., & Peñalver, E. A. (2021). Multimodal discourse in digital storytelling: An assessment tool proposal. *Computer-Assisted Language Learning Electronic Journal*, *22*(2), 14-25.

Vishnyakova, O. D., & Vishnyakova, E. A. (2022). Linguistic and cultural knowledge acquisition in terms of the multimodal approach to EIL studies. *Professional Discourse & Communication*, *4*(1), 81-92. https://doi.org/10.24833/2687-0126-2022-4-1-81-92

Waang, P. (2023). Maximizing the potential of multimedia in Indonesia: Enhancing engagement, accessibility, and learning outcomes. *Journal of Appropriate Technology*, *9*(3), 235-245. https://doi.org/10.37675/jat.2023.00409

Warschauer, M., & Whittaker, R. (2015). The Role of Digital Tools in the Language Learning Process. *Language Learning & Technology, 19*(2), 4–9. https://doi.org/10125/44464

Zhou, Y., & Wei, M. (2018). Strategies in technology-enhanced language learning. *Studies in Second Language Learning and Teaching*, *8*(2), 471-495.