TEACHING & LEARNING ENGLISH IN MULTICULTURAL CONTEXTS
http://jurnal.unsil.ac.id/index.php/tlemc/index
E-ISSN : 2541-6383
December 2025, Vol. 9 No. 2

DMC IN EFL CLASSROOM: ITS POSSIBLE PEDAGOGICAL VALUES TO DEVELOP STUDENTS' DIGITAL LITERACY AND CREATIVE EXPRESSION

Raynesa Noor Emiliasari[1], Emi Emilia[2], Budi Hermawan[3]
*raynesanoor@gmail.com
Universitas Majalengka[1], Universitas Pendidikan Indonesia[1,2,3], Indonesia

ABSTRACT

This study investigates the implementation of Digital Multimodal Composing (DMC) in an Indonesian EFL context, with particular attention to students' composing processes and their application of the Grammar of Visual Design (GVD). Adopting a qualitative case study design, the study was conducted over one academic semester and integrated the Genre-Based Approach (GBA) with multimodal instruction. The participants were twenty-eight undergraduate students enrolled in an English Education program who worked in four groups to produce four digital posters as the final outcome of the Independent Construction stage. All posters were designed using the Canva application and represented different genres, including narrative, informative, and persuasive texts. Data were collected through classroom video recordings, field notes, and semi-structured interviews with six purposively selected students. The findings show that students were able to integrate visual, verbal, and spatial modes in their poster designs, demonstrating an emerging understanding of representational, interactional, and compositional meanings. While students initially experienced difficulties in balancing modes and making deliberate design choices, scaffolded instruction through the GBA stages supported their gradual development of multimodal awareness. The posters illustrate students' ability to apply GVD principles in culturally relevant and socially meaningful contexts, highlighting the pedagogical value of DMC for fostering digital literacy, creative expression, and genre awareness in EFL learning.

Keywords: Digital Literacy, DMC, EFL Classroom, GBA

INTRODUCTION

The use of multimodal texts in foreign language learning has attracted increasing researchers and classroom practitioners' attention in recent years, particularly in relation to digital literacy development and creative expression in EFL classrooms. This growing interest builds on foundational work in multiliteracies and modality (Cope & Kalantzis, 2009; O'Halloran, 2004) and is further evidenced by recent studies on multimodal composing and digital meaning-making in language education (Hafner, 2019; Towndrow et al., 2022). Multimodality refers to integrating different modes of communication, such as verbal, visual, and audio texts, which complement each other to construct richer and more complex meanings (Kress & van Leeuwen, 2006). In English as a foreign language (EFL) pedagogy, this multimodal perspective underpins the use of multimodal texts to support students'

*Corresponding Author
Raynesa Noor Emiliasari
Email: raynesanoor@gmail.com

engagement with both verbal meaning and visual-audio elements inherent in contemporary communication practices (Hafner, 2019; Serafini, 2014; Towndrow et al., 2022). Thus, integrating multimodal texts into EFL learning can create a more dynamic and relevant learning experience while strengthening students' overall language competence.

Advances in digital technology have accelerated the transition to multimodal literacy as an essential component of language education. One innovative practice in this area is Digital Multimodal Composing (hereafter DMC), which is creating texts that combine multiple modes of communication—including written, visual, and audio elements—through digital platforms. DMC allows students to interactively utilize semiotic resources, integrating verbal, visual, and audio elements in language learning. This concept is rooted in the theoretical framework of multiliteracies, which emphasizes the importance of adaptability in constructing meaning through multiple modes of communication in the digital age (Cope & Kalantzis, 2009; The New London Group, 1996). With the increasing demand for multiliteracies in a globalized world, DMC is a strategic tool to support students in developing in-depth and contextual literacy skills.

In the EFL realm, DMC extends beyond language practice by enabling students to interpret and produce meaningful multimodal texts. Early studies on multimodal composition highlighted the pedagogical potential of integrating multiple semiotic modes in second language meaning-making (Yi, 2014), a perspective that has been further developed in more recent research on digital literacy and creative expression in language education (Jiang, 2017; Hafner, 2019; Towndrow et al., 2022). This process involves analyzing visual elements, such as representation, interaction, and composition, that collectively construct meaning in multimodal texts (Kress & van Leeuwen, 2006). In addition, through the application of visual theory in the production of digital multimodal texts, students gain a learning experience that is more relevant to the needs and challenges of the digital era.

While DMC integrates verbal, visual, and spatial elements and holds pedagogical potential for EFL learning, current research and classroom practices tend to prioritize learning outcomes over the composing process, resulting in limited understanding of how students operationalize visual theory in multimodal text production. Although Cho and Kim (2021) provide valuable insights into the impact of DMC on students' writing and learning outcomes, the present study shifts the analytical focus toward the composing process itself, foregrounding how students interpret visual theoretical concepts and translate them into the creation of meaningful multimodal texts. Accordingly, this study aims to investigate the implementation of DMC as a learning strategy that enables students to apply their understanding of the Grammar of Visual Design (Kress & van Leeuwen, 2006).

The Genre-Based Approach (hereafter GBA) is a relevant approach to support multimodal texts in EFL learning. This approach, as explained by Martin and Rose (2008), focuses on teaching texts within a broader genre framework, providing a structure for students to understand and produce multimodal texts that conform to the conventions of a particular genre. GBA allows students to recognize the features and conventions of different genres, which is crucial for analyzing and producing multimodal texts (Derewianka & Jones, 2012). Meanwhile, as discussed by Cope and Kalantzis (2009), DMC enables students to actively create texts that adhere to the genre conventions studied while integrating various modes of communication, such as written, visual, and audio elements.

In the context of EFL learning, applying GBA provides a clear structure for understanding various genres of texts, which can then be integrated with DMC practices. GBA helps students identify and analyze genre-specific features (Hyland, 2007), while DMC fosters an interactive approach to text creation relevant to contemporary digital communication (Jiang, 2017). The combination of these approaches supports contextualized language learning by promoting the application of language skills in various real-life scenarios, as outlined by the New London Group (1996) in their framework for multiliteracies.

A growing body of research has highlighted the pedagogical value of DMC in EFL classrooms, particularly in relation to students' writing outcomes and the development of digital literacy (e.g., Cho & Kim, 2021; Hafner, 2019; Towndrow et al., 2022). Other studies have also demonstrated how genre-based pedagogy supports students understanding of textual

organization and meaning-making across different genres (Emilia, 2012; Martin & Rose, 2007). Despite these contributions, little known about how Genre-Based Approach (GBA) can be meaningfully integrated with DMC in EFL contexts, especially with regard to the composing process itself. In particular, previous research has paid limited attention to how students interpret and apply visual theoretical concepts, such as Grammar of Visual Design (Kress & van Leeuwen, 2006), when producing multimodal texts within genre-oriented instruction. Addressing this gap, the present study explores the implementation across different genres of multimodal texts.

METHODS

This study employed a qualitative case study design to explore how Indonesian EFL students applied the GVD in their DMC practices. A case study approach was selected because it enables an in-depth examination of a bounded instructional context and its processes over time (Creswell, 2014; Yin, 2018; Merriam & Tisdell, 2016). Instead of examining learning outcomes, the study foregrounded students' composing processes and the ways visual theoretical concepts were applied in a genre-based instructional setting.

The participants were 28 undergraduate students enrolled in an English Education program at a public university in Indonesia. All students in the class participated in the study as part of a course-based instructional project. The students were between 18 and 23 years old, with 68% female and 32% male participants. All students voluntarily agreed to participate in the study and provided informed consent through a consent form distributed via WhatsApp prior to data collection.

Data were collected from three main sources: classroom recordings, field notes, and semi-structured interviews. Each data source served a distinct purpose in capturing different aspects of students' engagement with digital multimodal composing. Classroom recordings were used to document the implementation of GBA integrated with DMC during instructional sessions. The recordings focused on classroom interactions, teacher guidance, student discussions, and moments where students engaged with GVD concepts while designing multimodal texts. Field notes were taken by the researcher during each instructional session to capture students' behaviors, decision-making processes, and interactions that were not fully observable through recordings. Particular attention was given to how students discussed visual choices, negotiated meanings, and applied representational, interactive, and compositional principles of GVD during the composing process (Bogdan & Biklen, 2007). Semi-structured interviews were conducted to gain deeper insights into students' understanding of the GVD and their experiences in applying visual and verbal elements during digital multimodal composing. The interviews focused on students' interpretations of GVD concepts, challenges encountered throughout the composing process, and their reflections on the integration of GBA and DMC in the classroom (DiCicco-Bloom & Crabtree, 2006). From the total of 28 participants, six students were purposively selected for interviews to represent varying levels of engagement in multimodal composing. The interviews were conducted after students had completed their final multimodal projects to allow for reflective responses. Each interview lasted approximately 20–30 minutes and was conducted in Indonesian to facilitate more natural expression. All interviews were audio-recorded with participants' consent and subsequently transcribed for analysis.

Data collection was conducted throughout the implementation of a GBA-informed DMC project over one academic semester. Classroom observations and field notes were collected across all stages of the GBA, including Building the Field, Modelling of the Text, Joint Construction, and Independent Construction. These data sources documented how students engaged with the GVD principles during instructional activities and multimodal text production. The staged implementation of GBA provided a structured context for tracing students' composing processes, allowing the researcher to observe how visual, verbal, and spatial elements were introduced, scaffolded, and gradually applied by students in their multimodal texts.

The implementation of DMC in this study was guided by the stages of the Genre-Based Approach (GBA), which include Building the Field, Modeling of the Text, Joint Construction, and Independent Construction. These stages were adopted from the SFL-based genre pedagogy (Martin & Rose, 2007; Emilia, 2012) and adapted to accommodate the multimodal nature of digital text production. The adaptation involved integrating principles of the Grammar of Visual Design (GVD) into each stage to support students' multimodal meaning-making (Kress & van Leeuwen, 2006; Jewitt, 2009).

The building the field stage was conducted over two class meetings. During this stage, students were introduced to the basic concepts of the GVD, including representational, interactive, and compositional meanings (Kress & van Leeuwen, 2006). Guided discussions and small-group activities were used to help students explore how visual, verbal, and spatial elements function together in multimodal texts. Students analyzed selected digital texts and identified semiotic resources such as images, colors, typography, and layout commonly used in digital media. This stage aimed to build students' shared knowledge of multimodal meaning-making and prepare them for subsequent text analysis and production (Martin & Rose, 2007; Emilia, 2012).

The modeling of the text stage was implemented across two class meetings. Students were provided with exemplars of digital multimodal texts representing different genres. Through guided analysis, students examined how GVD principles were realized in authentic texts, focusing on the interaction between visual and verbal elements. The lecturer facilitated discussions that connected theoretical concepts of GVD with observable design choices in the texts. This stage enabled students to develop an explicit understanding of how multimodal meanings are structured and organized in digital genres, serving as a bridge between theory and practice (Jewitt, 2009; Serafini, 2014).

The joint construction stage took place over two meetings and involved collaborative multimodal text production. Students worked in small groups to design a draft multimodal text under the lecturer's guidance. During this stage, students applied GVD concepts collectively, negotiated design decisions, and received immediate feedback. The lecturer provided scaffolding through questioning, suggestions, and demonstrations to support students' emerging control of multimodal composing. This collaborative process allowed students to refine their understanding of multimodal design while gradually assuming greater responsibility for text construction (Emilia, 2012; Martin & Rose, 2007).

The independent construction stage was conducted over two class meetings and functioned as the culminating phase of the instructional cycle. Students independently produced a digital multimodal text as a final project, applying the visual, verbal, and spatial principles learned throughout the course. Students revised and refined their work based on prior feedback, focusing on coherence, clarity, and effective multimodal integration. This stage provided evidence of students' ability to operationalize GVD concepts autonomously within genre-oriented multimodal text production.

The data analysis was conducted iteratively and involved three primary data sources: classroom recordings, field notes, and semi-structured interviews. The analysis focused on understanding students' composing processes and their application of the GVD within the stages of the GBA. Classroom recordings were first transcribed verbatim and analyzed to examine how GBA stages were enacted during instruction and how students engaged with GVD concepts throughout the digital multimodal composing process. The analysis focused on classroom interactions, teacher scaffolding, peer discussions, and moments where students explicitly negotiated visual, verbal, and spatial design choices. These recordings allowed the researcher to trace how GVD principles were introduced, discussed, and gradually applied during multimodal text production. Field notes were analyzed alongside the recordings to enrich the interpretation of classroom events. The notes were used to document students' observable behaviors, decision-making processes, and challenges encountered during composing that were not always evident in audio recordings. Attention was given to students' problem-solving strategies, design revisions, and responses to feedback across different GBA stages. Semi-structured interview transcripts were analyzed using thematic analysis. An initial open coding

process was conducted to identify recurring themes related to students' understanding of GVD, their experiences in applying visual theory, and their reflections on multimodal composing. These codes were then grouped into broader themes, such as challenges in multimodal design, strategies for integrating visual and verbal elements, and perceived development of multimodal awareness. The interview data were used to complement classroom observations and to provide students' reflective perspectives on the composing process.

Findings from the three data sources were triangulated to enhance the credibility of the analysis by comparing patterns observed in classroom practices with students' reflections and documented behaviors.

FINDINGS AND DISCUSSION

Four groups created four posters. Each group used the Canva application to design the poster. Students created the posters as the final result of the DMC project, a product of independent construction. This poster illustrates their mastery of multimodal composing techniques learned during the deconstruction and joint construction processes. Before reaching the independent construction stage, students have gone through several revisions and discussions with the lecturer and analyzed poster examples to understand the techniques used.

Poster 1

The first group made the first poster. It was a narrative poster. The digital poster created is a multimodal representation that employs various modes, including visual, linguistic, spatial, and gestural. In general, this poster displays visual elements in the form of an illustration of a temple that is the main symbol of the legend of "Roro Jonggrang", a figure of a woman dressed in traditional Javanese clothes, and neutral and simple background elements. The verbal mode is shown through the text in the form of the title "Roro Jonggrang", the subtitle "Legend from Central Java", and a short narrative that explains the essence of the story. Meanwhile, the spatial mode is seen in the layout of the visual elements and text, which creates a hierarchy of information with the top highlighting the main information. At the same time, the bottom enriches the narrative with visual details and additional text. The gestural mode is represented by the female figure in the poster, standing with calm, formal hands in front of her body. This pose reflects elegance and politeness, key traits of Javanese culture, enhancing the traditional atmosphere of the Roro Jonggrang story.

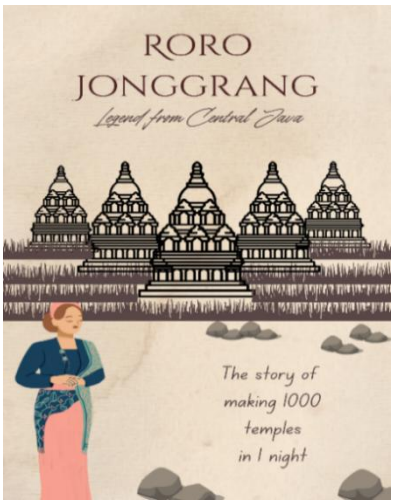

Figure 1. Poster 1

Compared to GVD of Kress and van Leeuwen, this poster reflects the effective use of representational, interactional, and compositional metafunctions. The legend of "Roro Jonggrang" is visually represented through local cultural symbols such as temples and traditional clothing, which strengthen the story's identity. In terms of interaction, there is no direct relationship with the audience. However, the woman's gesture creates a formal and polite impression, while the use of English in the text shows that this poster is intended for a wider audience. The poster's composition follows the principle of information value, where important elements such as the title and subtitle are placed at the top. At the same time, story details are presented at the bottom.

However, students faced challenges balancing visual and verbal elements and ensuring that all elements support the narrative without confusing the audience. Above all, this poster successfully demonstrates students' understanding of the principle of multimodality, especially in creating a strong cultural representation through minimalist visual elements. In the context of the learning objectives, this result shows that students have applied GVD theory well. However, there is room for further exploration, especially in strengthening interactive and gestural elements. This poster reflects success in understanding the theory and shows innovation in using local cultural elements as the primary identity.

Poster 2

The second poster made by group 2 brought up the fable as their topic for their project. The same is true with group 1; group 2 chose Canva as their media to design digital text. Based on the observation, this poster was made in the Joint and Independent Construction stages, where students discussed combining multimodal elements to create a straightforward narrative. In the Joint Construction stage, the collaborative approach helped the group share ideas and strengthen their understanding of arranging the elements into a coherent whole. The verbal dialogue and visual expressions used in the poster show cooperation in designing the elements.

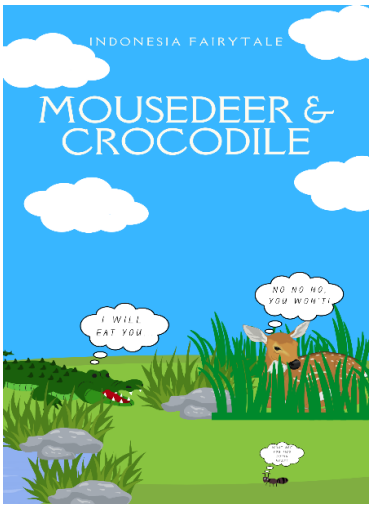

Figure 2. Poster 2

This poster is included in the narrative poster genre. This genre tells a story or event in visual and verbal form, often accompanied by illustrations or images that support the narrative. In this case, the poster aims to tell the Indonesian folktale, Mouse Deer & Crocodile, through three modes (verbal, visual, and spatial). This poster is intended to entertain and educate the audience, especially those unfamiliar with the Indonesian folktale, in an interesting and creative way.

The visual elements consist of illustrations of animals such as crocodiles, mouse deer, and ants, as well as a natural background of a river, grass, and a blue sky with white clouds. The verbal mode appears through the title text "Mouse Deer & Crocodile" which uses large

and straightforward typography at the top, as well as the character dialogues depicted with speech balloons: "I will eat you" by the crocodile and "No, no, no, you won't!" by the mouse deer. The spatial mode is seen in the organized layout, with key elements such as the title at the top to attract attention. At the same time, the illustrations and dialogues are placed in the middle to guide the audience through the narrative.

This poster's visual and verbal modes interact well, primarily through speech balloons that connect the dialogues to their respective characters, creating a direct relationship between the text and the images. The visual background supports the narrative by depicting the story's typical environment, namely the river and grass, which emphasize the location of the conflict between the mouse deer and the crocodile. The choice of bright colors, such as blue for the sky and green for the grass, creates a friendly atmosphere and attracts the attention of young audiences.

In terms of conformity with the GVD theory, this poster utilizes the representational metafunction through visual elements that depict the story scene in concrete terms: the crocodile in the water, the mouse deer on the riverbank, and the small ants at the bottom that add detail to the natural life. The interactional metafunction is seen in the dialogue in the speech balloons, which provide the audience with the character's perspective. However, there is no direct interaction (such as looking at the audience). From a compositional aspect, this poster follows the principle of salience by making the title and speech balloons the most striking elements. At the same time, the background remains supportive without distracting from the story.

The challenge students faced was ensuring that the visual elements were not too simple for older audiences and strengthening the relationship between the visual and verbal modes to be more expressive. However, the success of this poster lies in its ability to convey the story in a simple but effective way, using attractive colors and layouts. In the context of the learning objectives, this poster shows a good understanding of the principle of multimodality, especially in creating a balance between the visual and verbal to support the narrative of the story.

Poster 3

The poster created by group 3 shows three modes: verbal, spatial, and visual. The genre of the poster is an informative, persuasive poster. It is designed to inform and persuade the audience to take a specific action, namely getting vaccinated against COVID-19. This poster reflects work produced during the Independent Construction stage and the DMC process. The independent nature of the project is evident in the students' ability to combine modes without direct assistance from the lecturer. The students' choices of language, images, and layout show they have internalized the grammar of visual design theory and independently applied it in their final project.

Figure 3. Poster 3

The visual mode is shown through an illustration of a health worker administering a vaccine to a patient. Both figures wear masks, emphasizing compliance with health protocols. Other visual elements include the virus symbol in the letter "O" in the word "COVID-19" and curved lines in the corners that provide decorative elements and create visual dynamics. The verbal mode is present in the main text "COVID-19 Vaccination," the slogan "Protect everyone you love by getting vaccinated!" and the call to action "Don't be afraid to get vaccinated!" The spatial mode is seen in the organized layout, with a large title at the top as the most prominent element, followed by persuasive text in the middle, and a call to action at the bottom.

This poster reflects the representational metafunction in the GVD by showing a specific action, namely the vaccination process, which directly supports the meaning of the representation. The interactional metafunction is seen through the relationship created between the illustration and the audience, where the visualization of a person being vaccinated provides a positive and calming message. Although there is no direct eye contact with the audience, the atmosphere of the poster still creates a sense of security and support through the illustration and persuasive text. From a compositional aspect, this poster successfully utilizes the principle of salience by making the large title and illustration the most striking elements, while the other elements support the flow of information vertically from top to bottom.

The interaction between the visual and verbal modes in this poster is excellent. The illustration of the vaccination action reinforces the verbal message about the importance of protecting loved ones with vaccination. Students faced a challenge to add further visual appeal to increase audience attention, for example, through color variations or additional, more dynamic visual elements. However, this poster successfully conveys the main message clearly and persuasively, effectively integrating the visual and verbal modes. In the context of learning objectives, this work reflects a good understanding of the principle of multimodality, especially in composing relevant health messages, based on the theoretical approach of the Grammar of Visual Design.

Poster 4

The fourth poster is included in the persuasive educational poster genre. Its goal is to educate the audience, especially children and families, about the importance of consuming nutritious food for body health and to persuade them to implement a healthy diet as part of everyday life.

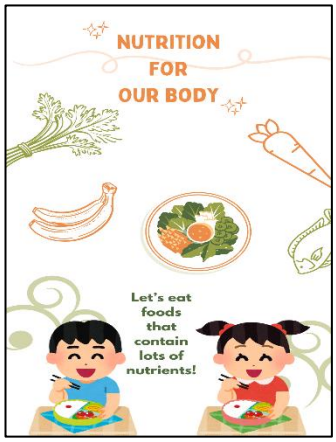

Figure 4. Group 4 poster

This poster is a multimodal representation that uses visual, verbal, and spatial modes to convey the importance of nutrition for the body. The visual mode is displayed through illustrations of various nutritious foods, such as vegetables, fruits, fish, and rice, and two children enjoying healthy food. In addition, decorative elements such as curved lines and small symbols reinforce the cheerful impression. The verbal mode is present in the title "Nutrition for

Our Body" and the slogan "Let's eat foods that contain lots of nutrients!" which provide direct information and invitations. The spatial mode is seen in the balanced arrangement of the poster elements, with the title at the top, food illustrations in the middle, and pictures of children with slogans at the bottom, creating an easy-to-follow flow of information.

From a Grammar of Visual Design perspective, this poster utilizes the representational metafunction by depicting a direct relationship between the healthy food illustrations and the message about nutrition. The interactional metafunction is seen in how the poster creates a relationship with the audience, especially children, through the image of two children smiling and enjoying nutritious food. Although there is no direct eye contact with the audience, their expressions project happiness, which can motivate the audience to imitate the behavior. Regarding compositional metafunction, the poster elements are arranged according to the principle of salience, where the title and the children's illustrations are the most prominent elements. In contrast, the food illustrations support the overall message.

Regarding the GBA and DMC process, the poster reflects the students' progression through the learning stages. During Building the Field, they likely explored nutritional concepts and existing public health campaigns. They analyzed how professional posters integrate visuals and text effectively at the Modeling the Text stage. Collaboration in joint construction allowed them to experiment with ideas, refine layouts, and receive feedback, evident in the balance of modes and thoughtful design.

Students' difficulties in designing digital multimodal texts

Analysis of classroom observations and interview data revealed three major themes concerning students' difficulties in designing digital multimodal texts. These themes highlight not only students' conceptual and practical challenges but also the pedagogical responses required to support multimodal literacy development in EFL contexts.

Limited Conceptual Understanding of Multimodality as Meaning-Making

The first and most fundamental difficulty identified in this study concerns students limited conceptual understanding of multimodality. Although students were familiar with various digital learning materials and technologies, most of them were initially unaware of the term *multimodal* and its theoretical implications. Classroom observations during the early stages of instruction showed that students tended to equate digital writing with typing text using digital tools, while visual and spatial elements were treated as secondary or decorative components rather than as meaning-making resources. This difficulty was clearly reflected in excerpt 1.

Excerpt 1

S2: "Before this class, I thought digital writing was just writing using a laptop or phone. I never thought about images or layout as part of meaning."

S4: "I often *use pictures and colors, but I didn't know that they have specific functions or names like visual mode or spatial mode."*

These responses in excerpt 1 indicate that students' engagement with digital media prior to the intervention was largely intuitive and experience-based, rather than grounded in explicit multimodal knowledge. This finding aligns with multimodality and multiliteracies research, which emphasizes that exposure to digital texts does not automatically result in multimodal competence (The New London Group, 1996; Cope & Kalantzis, 2015; Hafner, 2019). From a semiotic perspective, effective multimodal composition requires learners to develop *semiotic awareness*—the ability to recognize and purposefully integrate multiple modes to construct meaning (Kress & van Leeuwen, 2006; O'Halloran et al., 2015). Without such awareness, the communicative potential of multimodal texts remains underutilized.

Difficulty in Applying Multimodal Concepts and Metalanguage in Text Design

Beyond conceptual understanding, students also experienced difficulties in applying multimodal concepts during the composing process. Observation data revealed that students

often added images, colors, or layouts superficially, without considering how these elements contributed to meaning construction. Decisions related to framing, spatial organization, and visual salience were rarely justified unless explicitly prompted by the lecturer.

Interview data further confirmed this difficulty. Several students reported encountering multimodal terminology—such as *visual mode*, *verbal mode*, *spatial mode*, and *gestural mode*—for the first time in their English lessons as can be seen in Excerpt 2 below.

Excerpt 2

S1: First, I was confused when the lecturer talked about visual and spatial modes. I didn't know how to connect those concepts with my project."

S6: "Choosing images was easy but deciding why I chose them and how they matched the text was difficult."

These data suggest a clear theory–practice gap in students' multimodal composing. While students were comfortable using digital elements, they struggled to operationalize multimodal theory in purposeful design choices. Similar challenges have been documented in recent DMC studies, which report that learners often possess technical digital skills but lack guidance in composing multimodal texts strategically (Hafner & Ho, 2020; Towndrow et al., 2020; Jiang, 2018; Yeh, 2022). Jewitt et al. (2016) argue that multimodal competence develops through guided practice that explicitly connects theoretical concepts with concrete design actions—an insight that resonates strongly with the findings of this study.

Pedagogical Mediation through Simplified Multimodal Metalanguage

An important insight emerging from the data relates to the lecturer's pedagogical response to students' difficulties. Classroom observations showed that the lecturer deliberately simplified multimodal terminology to make abstract concepts more accessible. For example, mode was explained as *form*, *visual* as *what can be seen*, *verbal* as *words*, *spatial* as *placement*, and *gestural* as *body movements*. This strategy was consistently used during the early stages of instruction to support students' initial engagement with multimodal concepts.

Rather than indicating a limitation of instruction, this simplification functioned as pedagogical mediation. From a theoretical perspective, this approach aligns with Kress and van Leeuwen's (2006) argument that pedagogical metalanguage should be adapted to learners' levels of understanding. Serafini (2014) and Jewitt (2009) similarly emphasize that students' understanding of multimodal concepts develops gradually as theory is connected to practice through scaffolding and guided explanation.

This finding can also be understood through the framework proposed by Selfe and Selfe (2008), who argue that multimodal literacy involves technological, compositional, and social dimensions. In this study, students' prior experiences were largely limited to the technological dimension, while compositional and social understandings of multimodal meaning were underdeveloped. Recent studies likewise suggest that learners often require explicit mediation to transfer prior digital skills into academic composing contexts (Boche & Henning, 2015; DePalma & Alexander, 2015; Xue-quan & Zhang, 2020; Januarty & Ni'ma, 2018; Qi, 2023).

As instruction progressed, students began to perceive multimodal writing as a meaningful and creative alternative to traditional text-based writing. Interview data indicate that visual elements such as images, diagrams, and colors enabled students to express ideas that were difficult to convey through written text alone. This shift echoes findings from recent research highlighting the potential of multimodal composing to enhance engagement, creativity, and inclusivity in EFL learning (Qi, 2023; Starčič & Turk, 2016)

Overall, this theme highlights that students' difficulties in designing digital multimodal texts are not merely individual shortcomings, but developmental challenges that require deliberate pedagogical mediation, shared metalanguage, and scaffolded instruction.

CONCLUSION

The study highlights the potential of digital multimodal composing (DMC) as an effective literacy activity that improves students' ability to express themselves and be creative by involving them in producing multimodal texts. At the same time, this study demonstrates that EFL students' difficulties in designing digital multimodal texts extend beyond technical challenges and are primarily rooted in limited conceptual understanding of multimodality, difficulties in applying multimodal concepts during composing, and unfamiliarity with multimodal metalanguage. Although students were accustomed to using digital tools, they initially perceived digital writing as a technical activity rather than a meaning-making process that integrates multiple semiotic resources. The findings further show that pedagogical mediation, particularly through simplified metalanguage and scaffolded instruction within a Genre-Based Approach, played a crucial role in supporting students' gradual engagement with multimodal concepts and practices. Integrating Digital Multimodal Composing within a GBA framework enabled students to move from intuitive use of visual elements toward more deliberate and theoretically informed multimodal design. Overall, this study highlights the importance of process-oriented, theory-informed instruction in developing multimodal literacy and creative expression in EFL classrooms, contributing to current discussions on effective multimodal pedagogy beyond outcome-focused perspectives.

REFERENCES

Boche, B., & Henning, N. (2015). Multimodal scaffolding in a digital writing workshop. *Journal of Adolescent & Adult Literacy, 59*(1), 1–11.

Bogdan, R. C., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theory and methods* (5th ed.). Pearson.

Cho, H., & Kim, S. (2021). Exploring the use of digital multimodal composing in students' writing outcomes. *Journal of Digital Literacy Studies, 8*(3), 55–73.

Cope, B., & Kalantzis, M. (2009). "Multiliteracies": New literacies, new learning. *Pedagogies: An International Journal, 4*(3), 164–195. https://doi.org/10.1080/15544800903076044

Cope, B., & Kalantzis, M. (2015). *A pedagogy of multiliteracies: Learning by design*. Palgrave Macmillan.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE.

DePalma, M., & Alexander, K. (2015). A multimodal approach to literacy teaching: Insights from the classroom. *English Journal, 104*(2), 53–57.

Derewianka, B., & Jones, P. (2012). *Teaching language in context*. Melbourne: Oxford University Press.

DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical Education, 40*(4), 314–321. https://doi.org/10.1111/j.1365-2929.2006.02418.x

Emilia, E. (2012). *Genre-based approaches to writing instruction: An Indonesian perspective*. LAP Lambert Academic Publishing.

Hafner, C. A. (2019). Digital multimodal composing in second language learning. *Language Learning & Technology, 23*(2), 1–13.

Hafner, C. A., & Ho, M. (2020). Digital multimodal composing and academic literacy development. *System, 89*, 102190. https://doi.org/10.1016/j.system.2019.102190

Hyland, K. (2007). *Genre and second language writing*. University of Michigan Press.

Januarty, R., & Ni'ma, R. (2018). Multimodal literacy practices in EFL classrooms. *Indonesian Journal of Applied Linguistics, 8*(2), 353–362. https://doi.org/10.17509/ijal.v8i2.13303

Jewitt, C. (2009). *The Routledge handbook of multimodal analysis*. Routledge.

Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge.

Jiang, L. (2017). The potential of digital multimodal composing in second language development. *Educational Research and Reviews, 12*(4), 224–233. https://doi.org/10.5897/ERR2016.3059

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.

Martin, J. R., & Rose, D. (2007). *Working with discourse: Meaning beyond the clause*. Continuum.

Martin, J. R., & Rose, D. (2008). *Genre relations: Mapping culture*. Equinox.

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.

O'Halloran, K. L. (2004). *Multimodal discourse analysis: Systemic functional perspectives*. Continuum.

O'Halloran, K. L., Tan, S., Smith, B. A., & Podlasov, A. (2015). Multimodal meaning-making. *Visual Communication, 14*(4), 409–431. https://doi.org/10.1177/1470357215590738

Qi, Y. (2023). Digital multimodal composing and its impact on EFL learners' creativity. *English Language Teaching, 16*(3), 45–59. https://doi.org/10.5539/elt.v16n3p45

Serafini, F. (2014). *Reading the visual: An introduction to teaching multimodal literacy*. Teachers College Press.

Starčič, A. I., & Turk, Ž. (2016). Empowering students' engagement through multimodal compositions. *Education and Information Technologies, 21*(2), 361–374. https://doi.org/10.1007/s10639-014-9323-6

The New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review, 66*(1), 60–92. https://doi.org/10.17763/haer.66.1.17370n67v22j160u

Towndrow, P. A., Nelson, M. E., & Yusuf, W. F. (2020). Digital multimodal composing in English language classrooms. *TESOL Quarterly, 54*(4), 1063–1088. https://doi.org/10.1002/tesq.579

Towndrow, P. A., Nelson, M. E., & Yusuf, W. F. (2022). Multimodal literacy and digital composing practices. *Language Teaching Research, 26*(4), 597–617. https://doi.org/10.1177/1362168820911169

Xue-quan, Y., & Zhang, L. (2020). Scaffolding multimodal composition in EFL contexts. *Asian-Pacific Journal of Second and Foreign Language Education, 5*(3), 1–15. https://doi.org/10.1186/s40862-020-00092-5

Yeh, H.-C. (2022). Fostering multimodal literacy through digital composing. *Computer Assisted Language Learning, 35*(5–6), 1032–1055. https://doi.org/10.1080/09588221.2020.1825093

Yi, Y. (2014). Possibilities and challenges of multimodal literacy practices. *Journal of Adolescent & Adult Literacy, 57*(4), 296–306. https://doi.org/10.1002/jaal.250

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial support provided by the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia (Kemendikbudristek). The authors also extend their sincere appreciation to the students who voluntarily participated in this study and to the lecturer who supported the classroom implementation of the digital multimodal composing project. Finally, the authors thank the anonymous reviewers for their constructive feedback, which substantially improved the quality of this article.