JoAL
Journal of Applied Linguistics

JOURNAL OF APPLIED LINGUISTICS
Vol. 6 No. 1,July 2026
e- ISSN 2809 - 1922 , p-ISSN 2809 - 1914
Available online at:
https://journal.eltaorganization.org/index.php/joal/index

Analyzing Multimodal Text Use in Teaching Vocabulary: A Qualitative Study at SMPN 40 Medan

Nursella Subadar[1], Israya Oktaviana Sinaga[2], Dessy Natalia Sinaga[3]
Dolli Rotua Sinaga[4*], Jonris Tampubolon[5]
[1,2,3,4]PUI-PT ELT on Linguistics, Literature & Translation, Universitas Prima Indonesia, Indonesia
[5]English Education, Universitas Timor, Indonesia

Correspondence Email : dollirotuasinaga@unprimdn.ac.id

ARTICLE INFO

Article history:
Received February 6, 2026
Reviewed March 24, 2026
Revised April 2, 2026
Accepted April 3, 2026
Available online April 4, 2026

Keywords:

Multimodal texts; vocabulary acquisition; EFL instruction

Abstract

This study investigates the use of multimodal texts in teaching English vocabulary at SMPN 40 Medan, focusing on how teachers integrate multimodal resources, how students perceive these practices, and which types of multimodal texts are considered most helpful. Grounded in theories of multimodality and multimedia learning, the study addresses a gap in Indonesian junior high school contexts where vocabulary instruction often remains textbook-centered despite curriculum encouragement for digital literacy and student-centered learning. This study contributes to the limited empirical research on multimodal vocabulary instruction in Indonesian junior high school contexts. Using a qualitative descriptive design, the research draws on classroom observation, semi-structured interviews, and analysis of instructional materials to triangulate findings. The analysis shows that multimodal texts, particularly digital slides that combine text and visuals, support vocabulary learning by increasing engagement, clarifying meaning, and strengthening retention. Interactive quizzes enhance motivation and opportunities for review, while short audio and video materials support pronunciation and contextual understanding when guided through pausing and focused tasks. However, implementation is shaped by unstable internet access, limited devices, and the need for occasional translation support. Overall, the study highlights that multimodal vocabulary instruction is most effective when teachers connect visual, audio, and video input with meaningful practice and reinforcement while adapting to technological constraints.

INTRODUCTION

In recent years, multimodal learning has become increasingly central to English language education as students interact with texts, images, videos, and interactive media both inside and outside the classroom (Çakmak et al., 2021; Li et al., 2022). These developments reflect global educational trends that position multimodal input as an effective way to support vocabulary development, engagement, and meaning-making in second language learning.

In Indonesia, this orientation is reinforced by the Kurikulum Merdeka, which promotes student-centered learning, digital literacy, critical thinking, and exposure to varied text types. Multimodal texts,

Published by English Lecturers and Teachers Association (ELTA)

DOI: 10.52622/joal.v6i1.659

including images, videos, interactive platforms, and digital worksheets, align with these curriculum goals by helping students construct meaning through multiple semiotic resources.

However, empirical evidence on how multimodal texts are implemented in vocabulary instruction in Indonesian junior high schools remains limited. Although the curriculum encourages digital and multimodal learning, vocabulary teaching in many classrooms still depends heavily on textbooks and verbal explanation. This gap between curriculum expectations and classroom practice makes SMPN 40 Medan a relevant site for investigation. Preliminary observations suggest that teachers at the school have begun using digital and visual materials in English lessons, yet these practices have not been systematically documented.

Therefore, this study explores how multimodal texts are used in teaching vocabulary at SMPN 40 Medan, how teachers and students perceive these practices, and which kinds of multimodal materials are considered most helpful for vocabulary learning. By examining classroom implementation in a public junior high school context, the study seeks to contribute practical insights for vocabulary instruction and digital literacy development.

This study focuses on seventh-grade students at SMPN 40 Medan. It examines the multimodal resources used in classroom instruction, the ways teachers integrate those resources into vocabulary teaching, and students' responses to their use. The study does not attempt to evaluate all multimodal tools, but is limited to resources commonly used in the selected English learning context.

The findings of this research are expected to provide practical insights for English teachers regarding the integration of multimodal texts in vocabulary instruction. The study will also contribute to curriculum development by offering recommendations on how multimodal learning materials can support vocabulary acquisition in junior high schools. Additionally, this study may benefit future researchers by serving as a reference for investigations related to multimodal learning and vocabulary development. Policymakers and school administrators may also gain useful information to support digital literacy and modern learning approaches in English classes.

METHOD

This study employed a qualitative descriptive design with a sociolinguistic perspective to examine how multimodal texts were used in vocabulary instruction at SMPN 40 Medan (Kress, 2010; Suryanto & Fitrawan, 2023). The design was selected because it allows close description of classroom practice, participant perspectives, and the meanings attached to multimodal resources in a natural school setting.

The study was conducted in one eighth-grade English class at SMPN 40 Medan. The participants included one English teacher and one eighth-grade class involved in the observed vocabulary lesson. For interview data, one teacher and one student representative were selected purposively because they could provide direct information about the use of multimodal materials in the lesson.

Purposive sampling was used because the selected class had experience with multimodal materials such as PowerPoint slides, images, short videos, and quiz-based activities. This participant selection enabled the researcher to examine both pedagogical decisions and students' responses within the same instructional context.

Data were collected through classroom observation, semi-structured interviews, and document analysis. Classroom observation was used to record how multimodal texts were introduced, combined, and followed up during vocabulary teaching. Semi-structured interviews with the teacher and a student representative were conducted to explore perceptions, benefits, and constraints related to the use of multimodal texts. Instructional materials such as PowerPoint slides, images, short videos, and quiz activities were also analyzed to identify the forms of multimodal resources used in the lesson (Lestari et al., 2024).

Data analysis procedures

The qualitative data from observations, interviews, and instructional materials were analyzed thematically following repeated reading, open coding, category development, and theme generation (Jain et al., 2021). Observation notes, interview notes, and material analysis were compared across sources to identify recurring patterns related to teacher integration strategies, student perceptions, helpful multimodal resources, and implementation constraints.

Published by English Lecturers and Teachers Association (ELTA)

The main data set consisted of (a) one recorded vocabulary lesson conducted in a seventh-grade class and documented with an observation checklist; (b) interviews with one English teacher and one student representative; and (c) instructional materials, including a PowerPoint presentation with embedded images, a short video, and a quiz activity. These sources were triangulated to strengthen the credibility of the findings.

The thematic analysis was carried out through the following steps (Jain et al., 2021):

a. Familiarization: the researcher repeatedly reviewed the observation notes, interview notes, and instructional material analysis form.
b. Open coding: the researcher marked meaningful sections such as "visual clarity," "guided video viewing," and "offline workaround."
c. Axial coding: similar codes were grouped into categories linked to each research question, such as student perspectives, teacher integration strategies, and the most helpful multimodal resources.
d. Theme development: categories were organized into broader themes and checked across the different data sources.

To ensure trustworthiness, the researcher used methodological triangulation, maintained an audit trail through field notes and coded excerpts, and provided thick description of important classroom episodes. These procedures helped connect participants' statements with observed classroom practice and the instructional materials used during the lesson (Suryanto & Fitrawan, 2023).

FINDINGS

Students generally appreciate multimodal learning media, especially the combination of visuals and quizzes. One student said that visuals "make it easy and not boring," and that games support repetition and memory ("repeat many times"). These statements suggest that the texts of multiple modes motivate students and provide reinforcement (Researcher Student Interview Notes, 2026) which is consistent with the existing literature documenting the role of multimodal resources in fostering student engagement and providing scaffolding for vocabulary learning (Calafato & Simmonds, 2022; Li et al., 2022).

Some students also mentioned that videos can be "too fast," learning can be interrupted if the internet is slow, and some vocabulary becomes more confusing without translation support (Researcher Student Interview Notes, 2026). These concerns validate the existing realities of the classroom where the conditions surrounding access to and use of technology can impact the extent to which multiple modes are employed in learning.

Table 1. Student Perspectives: Benefits vs Problems

Aspect	Positive Responses	Problems Reported	Implication
Pictures/visuals	Easier, not boring; faster meaning	—	Use clear, familiar images for quick semantic access
Quiz games/apps	Fun; repetition supports memory	Internet slow; peer dependence for weaker students	Keep quiz accessible; mix group + individual checks; prepare offline option
Video/audio	Pronunciation + example sentences + context	Video too fast; loading issues; needs translation sometimes	Use short clips + subtitles + pause/replay + pre-task focus

Data from teacher interviews and observations indicate that the teacher embedded multimodal texts through a purposeful and structured sequence: stating and sharing the learning objectives and target vocabulary, using an image to build meaning, describing a brief video to provide contextual input, modeling the video to reinforce the listening/pronunciation, then moving to guided practice (sentence and match). This was followed with an interleaving quiz to review (Researcher Teacher Interview, 2026 and Teacher Observation Researcher, 2026).

This also exemplifies the type of integration that shows how multimodal learning can substantiate the principles of Dual Coding Theory and multimedia learning. Pairing a verbal description with a visual can serve to strengthen a cognitive memory trace and diminish the uncertainty principle. The short audio and video clips

served as a cue for the pronunciation and context of the target vocabulary (Mayer, 2005; Paivio, 1986). The teacher, in the observed lesson, made more than one explicit and direct reference to text and image (3/3 integration rating). The teacher also used numerous pictures to invite a meaning in the context of multimodal meaning. This was more than just the use of pictures for decorative purposes (Researcher Classroom Observation Sheet, 2026).

One teacher described several factors influencing integration: vocabulary concreteness (pictures for concrete words), the need for contextual/pronunciation support (video/audio), and practical limitations like poor internet and insufficient devices. The teacher mentioned using offline substitutes (pre-downloaded videos and pictures) for internet-free situations (Researcher Teacher Interview Notes, 2026). These instances support the rationale for Kurikulum Merdeka encouraging multimodal teaching, but infrastructural and school level constraints remain.

Table 2. Teacher Integration Strategy

Lesson Stage	Multimodal Resource	Teacher Action	Intended Vocabulary Support
1.Opening	PPT	States objectives + target vocabulary	Sets learning focus
2.Meaning-building	Images/icons	Elicits meaning; uses L1 equivalents when needed	Meaning clarity & concept mapping
3. Context input	Short video (~1 min)	Plays clip; pauses twice for "What word fits?"	Contextual meaning + attention control
4. Pronunciation	Audio model + drilling	Choral repetition; individual checks	Pronunciation accuracy
5.Guided practice	Picture-cards + writing task	Matching + sentence-making	Usage and consolidation
6. Review/check	Interactive quiz (10 items)	Runs quiz; reviews errors	Reinforcement + formative assessment

DISCUSSION

The teacher and students perceived the combination of text-based materials and visuals (PPT, pictures, and flashcards) as the most reliable and immediately useful multimodal resource. The teacher explained that text combined with visuals remained effective even under unstable internet conditions, while the student reported that pictures helped meaning become clear more quickly (Researcher Teacher Interview Notes, 2026; Researcher Student Interview Notes, 2026).

Students also responded positively to interactive online quizzes because the activities were engaging and allowed them to check their understanding. At the same time, both the teacher and the student noted that weaker learners sometimes depended on classmates during quiz activities. This pattern was also visible in the classroom observation and is consistent with studies showing that interactive resources can increase engagement but still require careful design to ensure broad participation (Calafato & Simmonds, 2022; Khasyi, 2024).

Many students stated that audio and video materials were helpful for learning pronunciation and seeing how words were used in sentences. However, they also reported falling behind when clips moved too quickly or when technical problems interrupted the lesson. In the observed class, the teacher addressed this issue by pausing the video, asking guiding questions, and replaying key segments. This strategy aligns with multimedia learning principles that emphasize segmentation and guided processing (Mayer, 2005).

Overall, the data indicate that multimodal texts aid vocabulary acquisition at SMPN 40 Medan by (1) boosting enthusiasm and building confidence through the use of visuals and interactive tasks, (2) enhancing understanding and retention through the linking of texts and visuals coupled with practice, and (3) in most cases, offering more than one benefit at a time: pictures for meaning, audio/video for pronunciation and context, and quizzes for reinforcement and assessment (Researcher Classroom Observation Sheet, 2026; Researcher Teacher Interview Notes, 2026; Researcher Student Interview Notes, 2026). These findings correspond with the aims of the proposal while strengthening the importance of examining the actual implementation in classrooms as opposed to the theoretical benefits

CONCLUSION

The present study found that vocabulary teaching at SMPN 40 Medan integrated both print and digital multimodal resources. The dominant mode was presentation text supported by visuals, especially PowerPoint slides and pictures, followed by short audio/video clips and interactive quizzes.

Students responded that pictures and quizzes make vocabulary learning simpler, and they appreciate the gamification of the learning process. However, they reported some difficulties with rapid video transitions, unreliable internet, and sometimes needing translation assistance.

The teacher integrated multimodal texts by sequencing the lesson from meaning-building to contextual input, pronunciation support, guided practice, and review. This pattern shows that multimodal resources are most useful when they are linked to lesson objectives and adapted to classroom constraints.

Text-based materials paired with visuals were viewed as the most reliable; interactive quizzes were the most motivating for practice and review; and audio/video materials were most helpful for pronunciation and contextual usage when they were short and carefully guided.

REFERENCES

Badhe, A., et al. (2025). Vocabulary development in junior high school learners: A Developmental perspective. Journal of Language Pedagogy, 14(2), 88–103.

Calafato, R., & Simmonds, R. (2022). Multimodal vocabulary learning and learner engagement. Language Learning & Technology, 26(3), 45–63.

Cárdenas-Claros, M., et al. (2023). Vocabulary learning through annotated and multimodal input. Applied Linguistics, 44(2), 243–265.

Çakmak, F., et al. (2021). Multimodal literacies in digital age classrooms. Journal of Digital Pedagogy, 6(4), 180–194.

Goud, J. (2025). Inductive qualitative analysis in classroom research. Research Methods Quarterly, 33(1), 11–28.

Ho, W., &Tai, K. (2020). Visual support in vocabulary retention. ELT Journal, 74(3), 247–257.

Jain, R., et al. (2021). Thematic analysis in qualitative educational research. Qualitative Insights, 18(2), 77–94.

Januarty, & Ni'ma. (2018). Vocabulary instruction strategies in Indonesian junior schools. ELT Perspective, 5(2), 55–70.

Khasyi, R. (2024). Student engagement with multimodal instruction. Indonesian Journal of EFL Teaching, 13(1), 10–25.

Kress, G. (2010). Multimodality: A social semiotic approach to contemporary communication. Routledge.

Kress, G., & van Leeuwen, T. (2001). Multimodal discourse: The modes and media of contemporary communication.

Laat, M., et al. (2020). Contextual analysis in multimodal classroom environments. Teaching & Learning Review, 28(3), 120–142.

Published by English Lecturers and Teachers Association (ELTA)

Lestari, D., et al. (2024). Classroom multimodal observation in EFL settings. Asia TEFL Journal, 21(1), 55–75.

Li, X., et al. (2022). Multimodal learning for better vocabulary acquisition. Second Language Research, 38(2), 230–255.

Mahendra, T., et al. (2025). Analysis of multimodal teaching materials in Indonesian schools. Journal of Instructional Texts, 5(1), 66–82.

Mayer, R. (2005). The Cambridge handbook of multimedia learning. Cambridge University Press.

MY, et al. (2024). Coding processes in multimodal qualitative research. Qualitative Research Forum, 19(1), 102–118.

Paivio, A. (1986). Mental representations: A dual coding approach. Oxford University Press.

Rahmanu, & Molnár. (2024). Multimodal practices enhancing language skills. International Journal of Literacy Studies, 9(2), 90–108.

Sari, M. (2024). English vocabulary teaching at Indonesian junior high schools. JELLT, 13(1), 88 100.

Sibanda, & Mhlanga. (2024). Content analysis in multimodal EFL materials. Instructional Text Analysis, 22(1), 15–29.

Suryanto, & Fitrawan. (2023). Qualitative descriptive design in educational research. Education Insight Journal, 10(4), 150–166.

Sutter, L., et al. (2024). Teacher perceptions of multimodal literacy. Journal of Media Literacy Education, 18(2), 24–39.

Ultavia, S., et al. (2023). Qualitative inquiry into EFL classrooms. Research on Language Pedagogy, 6(2), 120–138.

Published by English Lecturers and Teachers Association (ELTA)