

CHAPTER IV

FINDINGS AND DISCUSSION

A. Findings

This study presents the findings of research conducted at Muhammadiyah Elementary Schools 8 in Banjarmasin. Data were collected through classroom observations, semi-structured interviews with teachers and students, and documentation during the learning process. The observations were carried out across four classroom sessions: 27 April 2026 (topic: Ability), 4 May 2026 (Worksheet and Review), 11 May 2026 (topic: Holiday), and 25 May 2026 (Worksheet and Review). The findings are presented based on the two research questions in this study, namely: (1) the role of multimodal resources in supporting EFL students' vocabulary retention, and (2) the types of multimodal resources used by English teachers to improve EFL students' vocabulary retention.

1. The Role of Multimodal Resources in Supporting Vocabulary Retention

The following section presents the observation data collected across the four classroom sessions at SD Muhammadiyah 8 Banjarmasin. Each session is described in terms of the learning activities observed, student responses, and evidence of vocabulary retention.

Table 4.1 Summary of Classroom Observation

No	Date	Topic	Multimodal resources Used	Students Participant	Retention Indicators
1.	April 27, 2026	Ability	Gesture, speech, facial expression, writing	Active and enthusiastic	Recalled vocab, correct pronunciation, matched vocab with gesture.
2.	May 4, 2026	Worksheet and Review	Speech, writing, visual mode (picture)	Active and Focused	Recalled vocab to answer worksheet, correct pronunciation in review
3.	May 11, 2026	Holiday	Speech, music, auditory elements, visual mode (picture), writing, facial expression	Active during singing, less responsive in listening	Recalled vocab form sentences, wrote meaning on board
4.	May 25, 2026	Worksheet and Review	Speech, writing, visual mode (picture)	Active and Focused	Recalled vocab to answer worksheet, correct pronunciation in review

a. First Observation Session (April 27, 2026 - Topic: Ability)

During the first observation on 27 April 2026, the teacher introduced vocabulary related to the topic “Ability” using a combination of speech, gesture, facial expression, and writing on the whiteboard. The main activity was a movement-guessing game called “Guess the Gesture,” in which the teacher performed gestures representing vocabulary words and students identified and imitated them. The observation field notes recorded that all students were very enthusiastic and enjoyed learning English during this activity. Students paid close attention, participated actively, and responded eagerly. They imitated the teacher’s gestures and followed movement-based activities with considerable enthusiasm, while actively engaging with the teacher and peers. In terms of comprehension, the combination of gesture, facial expression, and whiteboard writing helped students grasp word meanings more concretely. By the end of the session, students demonstrated the ability to recall vocabulary words, pronounce them correctly, and match vocabulary with actions during the game, indicating early retention indicators from the multimodal instruction.

The teacher reflected on student responses to multimodal resources in the interview:

“Students usually respond positively to these resources. They enjoy learning through pictures and songs because the lessons become more engaging and enjoyable. They are also accustomed to writing new vocabulary in their notebooks, and many students are enthusiastic about adding new words to their ‘small dictionaries.’”(Teacher, interviewed on June 8, 2026)

b. Second Observation Session (May 4, 2026 - Worksheet and Review)

The second session on 4 May 2026, conducted approximately one week after the first session, focused on a worksheet covering the vocabulary material from the previous meeting. The teacher opened the session using speech to explain the instructions, and students completed written tasks including multiple-choice questions incorporating pictures and sentence construction based on pictures. During the review phase, the teacher corrected pronunciation verbally while students repeated. Observation field notes recorded that students were active and focused throughout. Significantly, students were able to recall vocabulary to answer worksheet questions and pronounce words correctly during the review,

indicating that the vocabulary introduced in the previous session had been successfully retained over the one-week gap.

c. Third Observation Session (May 11, 2026 - Topic: Holiday)

The third session on 11 May 2026 covered the topic “We Went Hiking Last Holiday” and was the most multimodally varied session observed. The teacher opened with an icebreaker and guided students to think about the lesson title before proceeding. Multiple modes were used in combination: speech, facial expression, music, auditory elements, visual mode, and writing. The teacher and students sang a song related to the topic together, which generated notable enthusiasm and active participation. Subsequently, students listened to audio while matching it with pictures in the textbook, an activity that engaged both auditory and visual processing simultaneously. However, some students were less responsive during this part because the audio was too fast and caused confusion.

Following the listening activity, the teacher guided students one by one to come to the front and write the meanings and pronunciations of vocabulary words on the whiteboard. Students demonstrated the ability to recall and write vocabulary meanings

correctly and were also able to construct sentences using the vocabulary. The teacher shared in the interview how multimodal resources support vocabulary learning:

“I consider pictures, songs, and vocabulary notebooks to be the most effective resources. Pictures help students connect words with meanings visually, songs help through repetition and rhythm, and vocabulary notebooks allow students to review words regularly. Since students use their notebooks continuously throughout the school year, the vocabulary becomes more familiar and easier to remember.” (Teacher, interviewed on June 8, 2026)

d. Fourth Observation Session (May 25, 2026 -Worksheet and Review)

The fourth session on 25 May 2026 followed the same format as the second session. Students completed a worksheet and participated in a vocabulary review, with the teacher using speech, writing, and visual mode through picture-based tasks. Students worked on multiple-choice questions and sentence construction tasks based on pictures. During the review, students pronounced

words correctly and recalled and matched vocabulary with pictures, again demonstrating retention of vocabulary from the previous instructional session approximately one week earlier.

The student interviewee expressed a positive overall response to the varied learning activities used across the sessions:

“ I really likes the various media and methods that have been taught because they don’t bore him.”

(Student, interviewed on June 8, 2026)

When asked which classroom activities best helped remember vocabulary, the student identified writing and memorizing vocabulary as the most effective, followed by pictures as very helpful, while noting that the listening activity was comparatively more difficult because the audio was too fast. The teacher confirmed that vocabulary retention is evaluated through classroom exercises, quizzes, speaking activities, and checking students’ vocabulary notebooks, and that students who can recognize, recall, and use vocabulary correctly in different contexts have successfully remembered the words.

Based on the observation data presented across all four sessions and the interview responses above, it can be concluded that multimodal

resources played three key roles in supporting vocabulary retention among the fifth-grade EFL students at SD Muhammadiyah 8 and 10 Banjarmasin. First, multimodal resources enhanced learner engagement, as evidenced by the consistently high levels of student enthusiasm and active participation observed across sessions, particularly during the gesture-based game and the singing activity. Second, multimodal resources supported meaning comprehension by enabling students to process word meanings through multiple sensory channels simultaneously, such as through the combination of gesture, pictures, writing, and speech. Third, multimodal resources provided memory cues and strengthened retention, as demonstrated by students' consistent ability to recall, pronounce, and apply vocabulary from previous sessions during the worksheet and review activities conducted one week later.

2. Types of Multimodal Resources Used by English Teachers to Enhance Vocabulary Retention

The observation data and interview responses identified several types of multimodal resources consistently used by the English teacher across the four classroom sessions. Based on the observation checklist and the teacher's

interview responses, the following types of multimodal resources were identified: speech, gesture (body language), auditory elements, verbal expression, facial expression, music, writing, and visual mode. The teacher also explained which resources she considered most effective in supporting vocabulary retention:

“I consider pictures, songs, and vocabulary notebooks to be the most effective resources. Pictures help students connect words with meanings visually, songs help through repetition and rhythm, and vocabulary notebooks allow students to review words regularly. Since students use their notebooks continuously throughout the school year, the vocabulary becomes more familiar and easier to remember.” (Teacher, interviewed on June 8, 2026)

This statement highlights three resource types that the teacher identified as most effective: visual resources (pictures), music (songs), and writing (vocabulary notebooks). The implementation of each resource type observed in the classroom is described below.

Table 4.2 Types of Multimodal Resources

No	Type of Multimodal Resources	Session 1	Session 2	Session 3	Session 4
1.	Speech	√	√	√	√
2.	Gesture	√	-	-	-
3.	Auditory Elements	-	-	√	-
No	Type of Multimodal Resources	Session 1	Session 2	Session 3	Session 4
4.	Facial Expression	√	-	√	-
5.	Music	-	-	√	-
6.	Writing	√	√	√	√
7.	Visual mode	-	√	√	√

Table 4.2 shows that seven types of multimodal resources were identified across the four observation sessions. Speech and writing appeared consistently in every session, while resources such as gesture, auditory elements, and music were used more selectively depending on the lesson topic. To provide a clearer picture of how each resource type was actually implemented in the classroom and how it supported students' vocabulary retention, each type is described in turn below.

a. Speech

Speech was used as the primary opening mode in all sessions that involved direct instruction. In the second session on 4 May 2026, the teacher began the lesson by explaining that the day's activity would involve working on a worksheet covering previously studied material, then explained how to answer the questions. Similarly, in the third session on 11 May 2026, the teacher opened with an icebreaker activity and used speech to guide students to think about the lesson title before transitioning to writing on the whiteboard. Speech was also consistently used during the review phase in all sessions, where the teacher verbally corrected pronunciation and explained vocabulary meanings. In the interview, the teacher explained that she uses verbal explanations when introducing vocabulary and that the combination of speaking and writing vocabulary words supports student understanding.

b. Body Language / Gesture



Figure 4.1 Students and teachers during "Guess the Gesture"

Gesture was the most prominently featured multimodal resource in the first session. The teacher used a movement-guessing game called “Guess the Gesture” specifically to help students remember vocabulary related to ability. Field notes confirmed that the teacher used this game for vocabulary memorization purposes, and students were observed following and imitating the teacher’s gestures and body movements during the activity. The gesture-based activity was the primary vocabulary retention strategy employed in that session, and it produced the highest level of student enthusiasm recorded across all observations. The teacher also recalled vocabulary

written on the whiteboard during the gesture game, combining the kinesthetic mode with the written mode.

c. Auditory Elements

Auditory elements in the form of recorded audio were used during the third session on 11 May 2026. Students were asked to listen to audio provided by the teacher and to match the audio content with corresponding pictures in the textbook, making this an audio-visual activity that engaged both auditory and visual processing simultaneously. However, the observation field notes recorded that some students were less enthusiastic during this exercise because the audio was too fast and caused confusion. In the student interview, the learner similarly noted that the listening activity was comparatively more difficult. This finding highlights that even when audio is paired with visual support, the pace and clarity of the audio itself are important factors in the success of the activity audio that is too fast can hinder comprehension regardless of the accompanying visual mode.

d. Facial Expression

Facial expression was used as a complementary resource during the first session on 27 April 2026. The field notes for the

types of multimodal resources used indicated that when playing “Guess the Gesture,” the teacher also used facial expressions alongside the gesture activity, using expressive facial cues to reinforce the meaning of vocabulary. Facial expressions were also used in the third session on 11 May 2026, where the observation record confirmed their use during the lesson. These facial cues functioned as additional non-verbal meaning-making resources that supported students in connecting emotional and contextual meanings to vocabulary words.

e. Music

Music was used in the third session on 11 May 2026, where the teacher and students sang a song relevant to the lesson material “We Went Hiking Last Holiday.” The observation field notes recorded that some students showed particular enthusiasm during the singing activity, and classroom interaction during this activity was notably positive. Crucially, the singing activity generated stronger student engagement than the audio recording-based listening exercise in the same session. This contrast between music (song with active participation) and audio recording (passive listening) suggests that the interactive and participatory nature of songs makes them more effective as multimodal resources for vocabulary retention than

audio recordings alone. The teacher confirmed this view in the interview, identifying songs as among the most effective resources because music helps through repetition and rhythm, and because the combination of hearing, vocalizing, and moving to a song creates a richer memory trace than passive listening

f. Writing



Figure 4.2 The students write vocabulary

Writing was the most consistently applied resource across all four observation sessions. In the first session, the teacher told students to write down vocabulary words on the blackboard in their own notebooks. In the third session, students went to the front one by one to write the meanings and pronunciations of vocabulary on the whiteboard. In the worksheet sessions (second and fourth), students engaged in written tasks including multiple-choice questions, sentence construction based

on pictures, and direction-based sentences. A particularly significant feature was the use of vocabulary notebooks as a long-term written resource: since the beginning of the school year, the teacher had encouraged students to maintain a special section in their notebooks called “Vocabulary,” which she referred to as their “small dictionary.” In the teacher’s interview, she described this practice as one of the most effective strategies, because students use their notebooks continuously throughout the school year and the vocabulary becomes more familiar and easier to remember with regular review. The student interview also affirmed the role of writing, with the learner identifying writing and then memorizing vocabulary as the most effective activity for remembering words.

g. Visual Mode

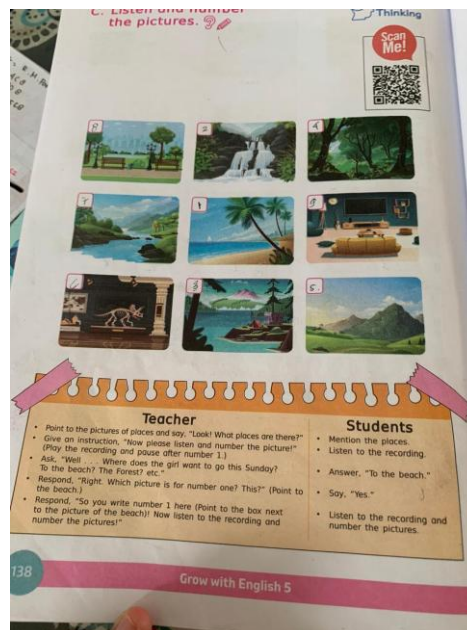


Figure 4.3 textbooks with illustrations that were used

Visual resources, primarily in the form of pictures from the students' textbook, were used in three of the four sessions. In the second session (4 May 2026), the worksheet included multiple-choice questions, some of which incorporated pictures, and students were asked to construct sentences based on those pictures. In the third session (11 May 2026), students matched pictures in the textbook with audio during the listening activity. In the fourth session (25 May 2026), students again worked with picture-based multiple-choice questions and sentence construction tasks. The teacher explained in the interview that

she usually uses pictures from the students' textbook when teaching vocabulary, as these visual resources help students connect words with meanings. She also acknowledged the occasional use of videos during special activities such as watching a movie together, though she noted that using videos is not very common because preparing a projector can be challenging and not all classrooms are equipped with one.

Overall, the observational data in Table 4.2 and the interview responses show interesting similarities as well as contrasts. Speech and writing were the only two resource types used in all four sessions, making them the most consistently applied resources in the classroom, while gesture, auditory elements, and music each appeared in only a single session and facial expression appeared in two sessions. However, when the teacher was asked directly which resources she considered most effective, she highlighted visual resources (pictures), music (songs), and writing (vocabulary notebooks) — not speech or gesture, despite speech being used in every session. This suggests that frequency of use in the classroom does not necessarily correspond to perceived effectiveness; a resource can function as a constant delivery tool (such as speech) without being regarded as the primary driver of vocabulary retention, while a less frequently used resource (such as music) can still be seen as highly memorable when it is used. The student interview data

reinforced part of this pattern, as the student identified writing and picture-based activities as the most helpful for remembering vocabulary, aligning with the teacher's emphasis on writing and visual resources, while also noting that the audio-based listening activity was comparatively harder to follow. Overall, the analysis of both observation and interview data indicates that writing and visual resources emerged as the two types most consistently supported by both classroom evidence and participant perceptions, while speech functioned primarily as a delivery mode that accompanied other resources rather than as a standalone driver of retention.

B. Discussion

This section discusses the findings in relation to the theoretical framework and previous studies presented in Chapter II. The discussion addresses both research questions in turn, examining how the classroom evidence aligns with or extends the existing literature on multimodal learning and vocabulary retention among elementary EFL learners.

1. The Role of Multimodal Resources in Supporting Vocabulary Retention

The three roles identified in the findings enhancing learner engagement, supporting meaning comprehension, and providing memory cues and strengthening retention are consistent with the theoretical framework established in Chapter II. The role of multimodal resources in enhancing learner engagement aligns with the view of Kress, (2010) and Lim et al., (2022), who argue that multimodal pedagogies involve the orchestration of different meaning-making resources to design learning experiences. The observation data confirm that activities integrating multiple modes in a participatory way, particularly the gesture-based game and the singing activity, consistently produced higher levels of student engagement than the audio-visual listening exercise, where students' enthusiasm was lower due to the speed of the audio. This is consistent with Kaminski, (2018) finding that young learners responded to multimodal input by joining in actions and imitating language, indicating a high level of involvement in the learning process.

The role of multimodal resources in supporting meaning comprehension is also supported by the data. As Mayer, (2002) explains, meaningful learning occurs when learners select relevant information from both verbal and visual channels, organize it, and integrate it with prior knowledge. The observation data showed that students who worked with picture-based tasks and gesture activities demonstrated better comprehension of word meanings. Even in the audio-visual listening activity where pictures were provided alongside audio, comprehension was hindered not by

the absence of visual support but by the speed of the audio, suggesting that the quality and pace of delivery within a multimodal activity matters as much as the combination of modes. The student's identification of pictures as resources that "can be seen right away and are easy to remember" aligns with Paivio, (2006) Dual Coding Theory, which holds that visual and verbal representations create interconnected memory traces, providing multiple retrieval routes.

The role of multimodal resources in providing memory cues and strengthening retention is the most directly supported by the delayed recall evidence observed in the worksheet sessions. The fact that students were able to correctly recall, pronounce, and apply vocabulary from the previous week's instruction during the worksheet and review sessions suggests that the multimodal delivery produced durable memory traces. This finding aligns with the conclusions of Aedo & Millafilo, (2022), who found that multimodal texts increased vocabulary acquisition and delayed retention among young EFL learners, and with Teng,(2023)) finding that richer multimedia input conditions produced stronger vocabulary learning and long-term retention. The contrast between the strong retention indicators following gesture and song activities compared to the lower engagement observed during the audio-visual listening exercise further supports Zhang & Zhang, (2024) finding that multimodal input is most effective when it actively engages learners across multiple channels in a participatory way, not merely by presenting information through more than one mode simultaneously.

2. Types of Multimodal Resources Used by English Teachers to Enhance Vocabulary Retention

The types of multimodal resources identified in this study speech, gesture, auditory elements, facial expression, music, writing, and visual mode are broadly consistent with the classification proposed by Apriani et al.,(2021), who identified similar categories in their study of EFL teachers in Indonesia. However, the present study adds specificity to this framework by examining how each resource type functioned in practice and how students responded to each type.

From all the learning resources observed, written and visual learning resources were the most consistently used across all sessions and produced the most consistent evidence regarding vocabulary retention. The vocabulary notebook system, in particular, represents a strategic implementation of writing as an ongoing multimodal resource rather than a one-time activity. The teacher's description of this "small dictionary" practice aligns with Schmitt, (2008) argument that written exposure helps learners recognize the spelling and written form of words, which is important for reading and writing accurately. Webb et al. (2023) similarly emphasize that writing activities encourage students to think more deeply about word meanings and usage, helping learners move from recognizing words to using them actively.

The gesture-based activity (“Guess the Gesture”) produced the highest level of student engagement across all observations and generated strong indicators of vocabulary retention. This finding aligns with Ismail et al., (2022), who found that the use of gestures in teaching verb vocabulary to young EFL learners improved students’ vocabulary understanding and helped them remember word meanings for a longer time, and with Oppici et al., (2023), who found that both performing and observing gestures have positive effects on foreign language vocabulary learning by creating stronger and more lasting memory traces.

Music in the form of a participatory song was notably more effective than audio recordings as a standalone resource. The singing activity in the third session generated stronger enthusiasm and interaction than the audio-based listening task in the same session. This contrast is consistent with Engh, (2013) finding that music supports vocabulary retention by activating both linguistic and affective memory systems, and with Fonseca-Mora et al., (2015) finding that combining musical input with other instruction resulted in significantly stronger vocabulary outcomes. The participatory and emotionally engaging nature of singing appears to create richer memory traces compared to passive listening to audio recordings.

A key finding emerging from the discussion of resource types is that the simultaneous integration of multiple modes consistently produced stronger vocabulary retention outcomes than reliance on any single mode. Sessions where the teacher combined speech, writing, visual resources, and either gesture or music

showed the highest levels of student engagement and the most robust indicators of retention. Audio recordings used alone produced the weakest response. This finding is consistent with the theoretical position of Mayer, (2024) and Paivio, (2006), both of whom argue that learning is most effective when information is processed through multiple complementary channels simultaneously. It also aligns with the conclusions of J. E. Liu & Lim, (2024), who note that teachers' conceptualizations of vocabulary teaching through various modes are central to effective multimodal pedagogy.

Bringing these two sets of findings together, the results show a clear connection between the three roles of multimodal resources identified earlier and the specific multimodal resource types applied in the classroom. The role of increasing learner engagement was most concretely realized through gesture and music, as seen in the strong enthusiasm generated by the "Guess the Gesture" game and the participatory singing activity, both of which actively involved students rather than positioning them as passive recipients of input. The role of supporting meaning comprehension was primarily carried out through visual resources and gesture, which allowed students to connect new vocabulary with concrete images and physical actions rather than relying on verbal explanation alone. Meanwhile, the role of providing memory cues and strengthening retention was most consistently supported by writing, particularly through the ongoing use of vocabulary notebooks, and by visual resources used repeatedly across sessions.

This mapping suggests that in the context of this study, no single type of multimodal resource fulfilled all three roles equally; rather, each role tended to be enacted more strongly by particular types of application, and the strongest overall retention outcomes emerged when several of these types especially writing, visual resources, and either gesture or music were combined within the same session.