

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

Heng Zhang

De La Salle University- Br. Andrew Gonzalez FSC College of Education (BAGCED) -M TEACHING ENGLISH LANG
(NT)(NTMTEL)

ABSTRACT: This action research explores the effectiveness of multimodal online instruction in supporting vocabulary acquisition among Chinese fifth-grade EFL learners. To address the limitations of traditional rote-based vocabulary teaching, the researcher implemented an eight-session extracurricular English course that integrated visual, auditory, and interactive modalities. Tasks included image matching, video clips, role-playing, digital games, and drawing-based activities.

A mixed-methods design was used to collect data through vocabulary pre- and post-tests, weekly formative homework assignments, student journals, classroom observations, and semi-structured interviews. The participants were 12 public school students aged 10–12 with varying levels of English proficiency. Results showed an average vocabulary score improvement of 11.58 percentage points, with the most notable gains among mid-level learners. Additionally, students displayed significantly higher engagement during game-based and image-supported tasks.

Qualitative data revealed that most students experienced positive emotions and greater confidence during multimodal lessons, particularly when learning was framed as interactive and playful. While some students initially struggled with technology, their motivation increased over time.

This study highlights the pedagogical potential of multimodal strategies in online EFL environments and emphasizes the importance of emotional engagement and task design in vocabulary learning. It provides practical insights for language teachers working with young learners in digital or hybrid settings.

KEYWORDS: Multimodal learning, vocabulary acquisition, online EFL instruction, gamification, elementary learners, action research

INTRODUCTION

In recent years, the role of multimodal learning in language education has gained increasing attention, especially in contexts where English is taught as a foreign language (EFL). In China, English is a compulsory subject beginning in the third grade of public primary education. However, vocabulary instruction at this level still heavily relies on rote memorization, textbook drills, and traditional teacher-centered delivery. While this approach may help students achieve short-term recall, it often results in low engagement, weak retention, and limited motivation for communicative use.

With the growing integration of educational technology—accelerated by the shift to online learning during and after the COVID-19 pandemic—educators have begun exploring new strategies to support vocabulary development among young learners. Multimodal instruction, which integrates text, audio, video, images, movement, and interaction, offers an alternative that is not only engaging but also cognitively and emotionally supportive.

This action research project was designed in response to an observed pedagogical problem: fifth-grade EFL students in extracurricular English classes showed low motivation and limited vocabulary retention when exposed to traditional teaching methods. By systematically designing, implementing, observing, and reflecting on a multimodal online intervention, the study aims to provide both practical solutions and theoretical insights. The focus is not only on vocabulary acquisition outcomes but also on understanding students' learning experiences, in line with the cyclical, reflective nature of action research.

BACKGROUND OF THE STUDY

In China's public school system, English is a core subject; however, the approach remains largely exam-oriented. Vocabulary teaching is typically conducted through repetition, copying, and isolated memorization, which often fails to promote

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

meaningful use or lasting retention. Learners frequently forget new words shortly after tests and display low enthusiasm in class. As an after-school English instructor, the researcher noticed that students became significantly more engaged when lessons involved multimodal elements—such as images, videos, and interactive games—instead of traditional textbook-based activities. These classroom observations served as the starting point for the action research cycle, leading to the identification of a specific issue: How can vocabulary learning be improved through a more engaging and sustainable method?

Importantly, while prior research has examined multimodal approaches in various educational contexts, few studies have focused on young EFL learners in online extracurricular settings in China. This study, therefore, addresses both a practical classroom challenge and a research gap by systematically testing and evaluating a multimodal teaching intervention through an action research framework.

RESEARCH QUESTIONS

Guided by the iterative and reflective process of action research, the study addresses the following questions:

1. How does multimodal online instruction promote vocabulary acquisition among fifth-grade Chinese EFL learners?
2. What are students' experiences and emotional responses to multimodal learning in an online extracurricular English course?

LITERATURE REVIEW

Multimodal Teaching and Vocabulary Retention

The theoretical foundation of this study draws heavily from multimedia learning theory (Mayer, 2009), which posits that learners understand and retain information more effectively when content is delivered through both auditory and visual channels. This dual-channel approach underpins multimodal teaching practices, especially in second language acquisition (SLA).

Zhang and Zou (2021), in a study with Chinese primary EFL learners, demonstrated that students exposed to multimodal materials—such as animated videos and annotated images—showed significantly better vocabulary retention in both immediate and delayed assessments compared to those taught using text alone. Zhao and Wang (2020) further emphasized that visual-contextual support helps learners form meaningful associations with unfamiliar vocabulary, reducing cognitive load and supporting lower-level learners in particular.

From an action research perspective, these findings inform the design of the intervention, as they highlight specific multimodal strategies likely to produce measurable improvements. This aligns with the researcher's goal to implement, observe, and refine instructional practices in response to learner needs.

In addition, qualitative studies (e.g., Chen & Huang, 2021) reveal that learners feel more engaged and emotionally connected to language when lessons include visual storytelling and audio-visual aids. These affective factors, often overlooked in traditional instruction, are central to the action research process, which values learner feedback and emotional engagement as key indicators for pedagogical improvement.

Gamified Learning and EFL Motivation

Gamification—defined by Deterding et al. (2011) as the use of game elements in non-game contexts—has proven effective in increasing student motivation and participation, especially in vocabulary learning. In the action research cycle, gamification serves as a practical innovation introduced to address the lack of learner motivation observed during the diagnostic phase.

Zarzycka-Piskorz (2016) found that students were more attentive and motivated when using platforms like Kahoot for vocabulary practice, benefiting from immediate feedback and the competitive, engaging format. Similarly, Liu and Li (2022) observed that game-based activities encouraged more spontaneous language use and improved speaking confidence in young EFL learners. Plass et al. (2015) emphasized that affective factors play a crucial role in gamified learning. This is especially true for language learners, as emotional engagement can effectively lower the affective filter, thereby facilitating better learning outcomes. Krashen (1982) proposed the Affective Filter Hypothesis, which suggests that learners' emotional states directly influence their success in language acquisition. If students feel anxious or uneasy, their learning effectiveness may be inhibited; conversely, when their emotional needs are positively supported, they are more likely to engage with and internalize new language material.

By integrating the perspectives of Plass et al. (2015) and Krashen (1982), it becomes clear that emotional responses not only impact students' engagement but also indirectly affect their learning outcomes, particularly in vocabulary acquisition. In online learning environments, affective factors are even more significant, as the absence of face-to-face interaction may lead to a lack of emotional involvement. Therefore, it is essential to incorporate affective support elements into task design—especially in the context of multimodal vocabulary learning.

Gamified strategies align well with learner-centered action research, allowing the teacher-researcher to continuously assess how changes in instructional design affect student behavior and attitudes. By incorporating elements like points, badges, and interactive games, this study builds on the idea that motivation and memory are interlinked, especially for younger learners with shorter

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

attention spans.

Moreover, in online learning environments, gamification plays a key role in sustaining interaction and feedback loops, which are essential for maintaining learner focus in remote contexts. These insights are directly applied to the development of lesson plans in this study's intervention phase and evaluated through student performance and feedback.

METHODOLOGY

Participants

The participants in this study were 12 fifth-grade students (6 boys and 6 girls), aged 10 to 12, from a public primary school in eastern China. All students were learning English as a foreign language, which is introduced in the formal school curriculum starting in Grade 3. The students in this study, however, attended an extracurricular online English course conducted after school hours by the researcher.

The group represented a mix of language proficiency levels—low, average, and high—based on their past academic performance and teacher observation. While they had some prior exposure to online platforms during the pandemic, their familiarity with interactive tools such as Kahoot or online games varied. All students and their guardians provided informed consent to participate in this action research.

Instruments

The following tools were employed to evaluate both vocabulary improvement and learner experience:

Vocabulary Tests (Pre-test and Post-tests): Designed by the instructor, these included multiple-choice picture identification, simple listening comprehension, and short fill-in-the-blank items. The tests focused on 40 key vocabulary words selected from the eight lessons.

Formative Assessments (Homework): Instead of conducting in-class formative tests, students completed short assignments at home after each lesson. These tasks included picture-word matching, sentence creation using target words, and vocabulary puzzles. This method was chosen to save in-class time for multimodal instruction.

Observation Checklist: The teacher maintained detailed observational notes on student participation, responsiveness, and emotional reactions during different types of tasks (e.g., video watching, role-play, game interaction).

Student Journals: Students wrote brief reflections after each class, often responding to teacher prompts about what they learned, which activity they liked most, and whether they encountered any difficulties. These were later analyzed for emotional tone and learning attitudes.

Semi-structured Interviews: Four students representing different engagement and proficiency levels were selected for individual interviews. Questions aimed to uncover their preferences, challenges, and suggestions for future lessons.

During the process of action research, these instruments served not only as means of data collection but also as tools for generating feedback at each stage of instruction. By analyzing students' vocabulary quiz results, observation records, and reflective journals, I was able to identify their learning needs and challenges based on both the collected data and my evolving understanding of the learners. Informed by these insights, I adjusted subsequent teaching strategies accordingly. These instruments constituted a critical component of the reflective cycle in action research, enabling the ongoing refinement of instructional design.

Procedure

In the planning stage of the study, a four-week multimodal intervention program was designed based on an analysis of students' backgrounds and identified learning needs. The primary objectives of the program were to improve students' vocabulary acquisition and oral communication skills. To support these aims, the instructional design incorporated a range of interactive and multimodal learning activities, including videos, animated slide presentations, Kahoot-based games, and role-playing tasks. Additionally, pre- and post-tests, along with homework assignments, were administered to evaluate students' vocabulary development throughout the intervention. The intervention was implemented over four weeks through an extracurricular online course held after regular school hours. Students participated in eight sessions, each approximately 40 minutes long, conducted twice a week using a video conferencing platform. The lessons followed a consistent pedagogical flow:

Vocabulary Input – through short video clips, animated slides, or illustrated flashcards;

Interactive Practice – using tools like Kahoot, digital games, or visual matching exercises;

Communication Task – role-play scenarios, drawing and describing, or TPR-based actions;

Post-lesson Homework – short assignments focusing on vocabulary consolidation.

Unlike traditional classroom instruction, time constraints during each session made it difficult to conduct quizzes in class. Therefore, formative assessments were administered as take-home tasks to measure students' understanding after each session.

During the action phase, eight online lessons were delivered using a variety of multimodal activities such as short videos, interactive slides, and role-playing tasks to facilitate students' vocabulary acquisition. Each lesson followed a structured sequence

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

that included vocabulary input, interactive practice, communicative tasks, and follow-up assignments. Students actively engaged in tasks such as Kahoot games and role-plays, with the lesson content specifically designed to enhance their vocabulary retention and spoken language proficiency.

The following table provides an overview of the teaching sequence:

Date	Lesson Theme
Feb.18-20	Giving out of consent forms Pre-test
Feb.22-March 1	Intervention1-4 Unit1: Family Members Unit2: Body Vocabulary Unit3: Food Unit4: Routines
March 3	Interview
March 4-11	Intervention 5-8
March 13	Interview Post test
March 18	Presentation of initial results
March 31	Submission of final paper

After each round of reflection, instructional content and methods were adjusted based on observational data and student feedback, forming a continuous cycle of improvement. Each reflective cycle contributed to the ongoing refinement of the lessons, ensuring they better addressed students' learning needs and enhanced their vocabulary acquisition. This iterative process not only facilitated student progress but also contributed to the ongoing development of my own teaching practice.

DATA ANALYSIS

This study uses a mixed-methods approach for data analysis, integrating both quantitative and qualitative data to comprehensively evaluate the impact of multimodal online teaching on students' vocabulary acquisition and emotional responses.

Quantitative Data (RQ1): Vocabulary test scores (pre-test and post-test) were analyzed using descriptive statistics, including the average improvement and individual score gains. Homework submission data was also reviewed to track students' learning progress and error patterns. Through the quantitative data, we were able to measure students' actual progress in vocabulary mastery, while also providing baseline data for subsequent qualitative analysis.

Qualitative Data (RQ2): Student journals, classroom observation notes, and interview recordings were coded thematically. Students' emotional responses were categorized as positive, neutral, or negative. By analyzing students' task preferences, classroom participation, and reported difficulties, several patterns were identified. The qualitative data provided a deeper insight into students' participation in classroom activities, revealing their emotional fluctuations during the learning process and responses to multimodal tasks.

Integration of Quantitative and Qualitative Data: To provide a more comprehensive analysis, we integrated the quantitative and qualitative data. Quantitative data (e.g., vocabulary test scores) provided an objective basis for assessing students' learning progress, while qualitative data (e.g., student journals, classroom observations, and interview feedback) revealed students' emotional responses, task preferences, and challenges in participating in multimodal tasks. By integrating these data, we were able not only to assess students' vocabulary learning outcomes but also to gain a deeper understanding of their emotional changes and engagement, uncovering the underlying reasons behind the teaching intervention.

Triangulation: To enhance the trustworthiness of the findings, we cross-checked multiple data sources. For instance, we compared students' journal responses with classroom observations and interview feedback to validate students' performance and emotional responses during different tasks. This process ensured the reliability and accuracy of the data analysis.

By combining quantitative and qualitative data, this study was able to comprehensively evaluate the impact of multimodal online teaching on students' vocabulary acquisition and reveal the potential effects of students' emotional responses and classroom participation on learning outcomes.

RESULTS AND DISCUSSION

This section presents the findings of the study based on both quantitative and qualitative data, and interprets them in light of the research questions and related literature. The results are discussed holistically, integrating test outcomes, formative assessments, classroom observations, student journals, and interview data.

Vocabulary Development Through Multimodal Learning

The vocabulary pre- and post-tests revealed notable improvement among participants. The average pre-test score was 55.17%,

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

while the post-test score rose to 66.75%, reflecting an average increase of 11.58 percentage points. This suggests that multimodal instruction had a positive impact on vocabulary acquisition.

When results were disaggregated by student proficiency level:

High-level learners (n = 3): Maintained consistently strong performance (80%+ on both tests), showing stable mastery and fluent usage of new vocabulary during speaking tasks.

Mid-level learners (n = 5): Showed the most significant improvement (10–18% score increases). These students responded well to visual support and game-based tasks.

Low-level learners (n = 4): Demonstrated modest gains (3–7%), but classroom engagement improved noticeably. They participated more in interactive tasks, particularly those involving visuals and drawing.

Figure 1 illustrates the improvement in test scores by proficiency level. The findings align with Zhang and Zou (2021), who found that multimodal input enhances word retention, especially when combined with interactive activities.

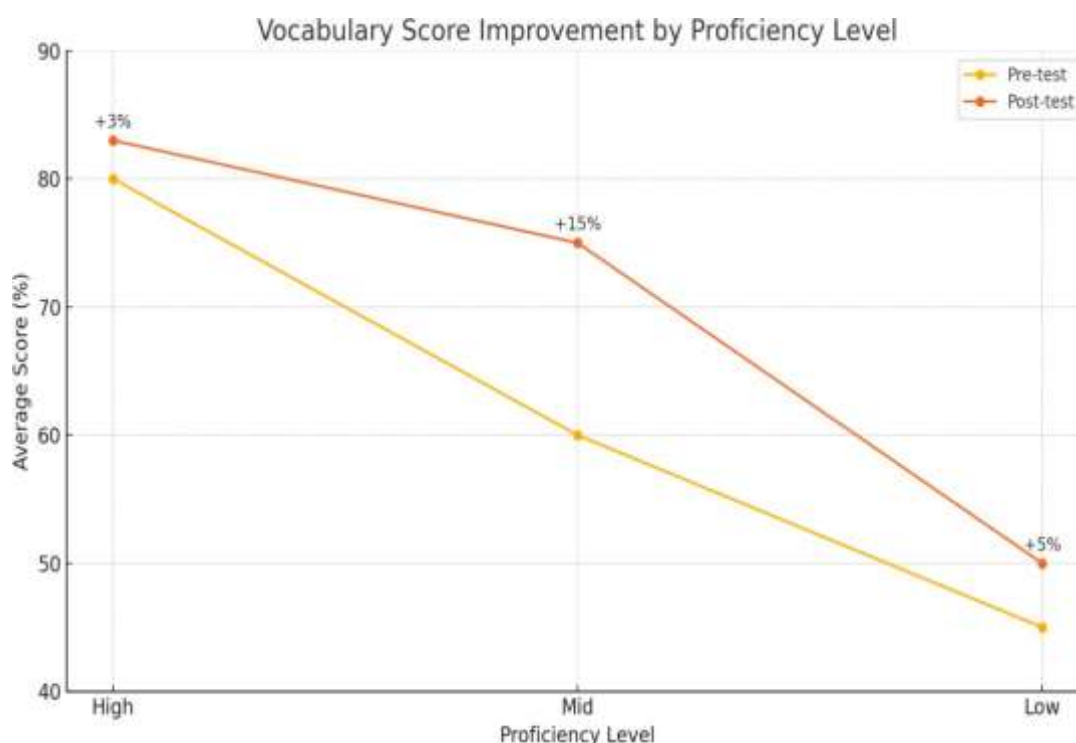


Figure 1: Vocabulary Score Improvement by Proficiency Level

The significant improvement among mid-level learners, in particular, suggests that the multimodal teaching approach, integrating visual, auditory, and kinesthetic elements, effectively supports vocabulary acquisition. This is consistent with Zhao and Wang's (2020) assertion that visual and contextual learning significantly aid vocabulary acquisition.

Although test scores fluctuated among individuals, the general upward trend supports the effectiveness of the multimodal online teaching model.

FORMATIVE TEST ANALYSIS

Formative assessments were conducted after each lesson to track students' vocabulary development. The average scores for each lesson indicate a steady improvement, with the highest score recorded in Lesson 8 (68.00%).

Lessons 1-2: Initial unfamiliarity with digital tools led to lower performance, with average scores around 58-60%.

Lessons 3-6: As students became familiar with the platforms and tasks, their performance improved, with scores averaging 62-64%.

Lessons 7-8: By the final lessons, students showed significant progress, with Lesson 8 achieving an average score of 68.00%.

Figure 2 below visualizes the trend in formative test scores across the intervention

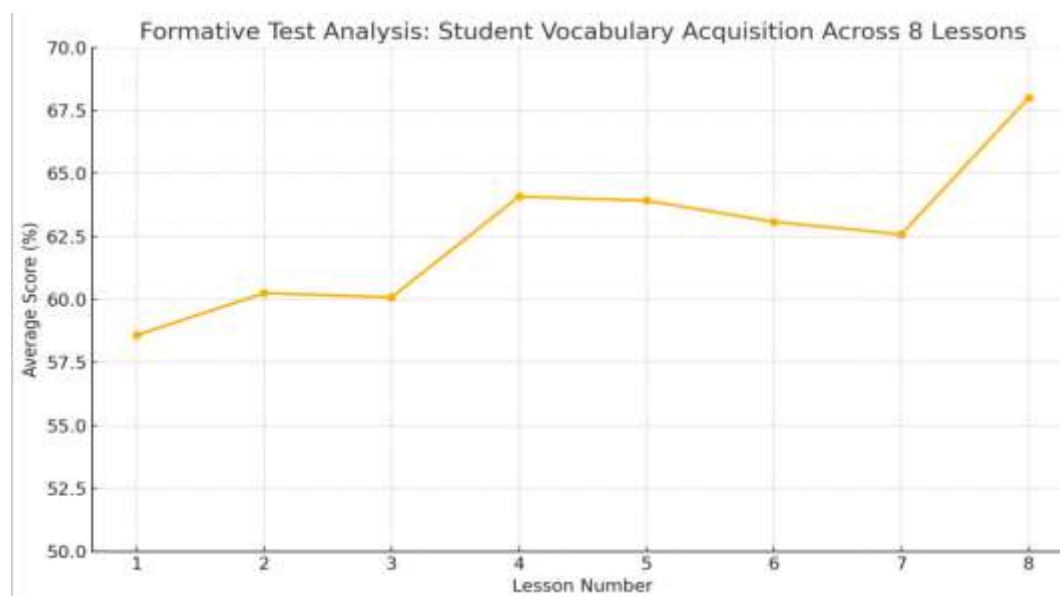


Figure 2: Formative Test Analysis: Students Vocabulary Acquisition Across 8 Interventions

The gradual increase in scores reflects the positive effect of the multimodal intervention in supporting vocabulary development. As students became more comfortable with the digital tools, their engagement and vocabulary retention improved. This trend aligns with Liu and Li (2022), who found that gamified learning increases student motivation and engagement.

STUDENT ENGAGEMENT AND EMOTIONAL RESPONSES

The analysis of students' emotional and behavioral responses to the multimodal learning intervention revealed significant variations in how these emotional responses impacted their participation and vocabulary acquisition. The students's emotional reactions were categorized into three groups: positive, neutral, and negative, with each group showing different patterns in their learning progress.

This action research explores the effectiveness of multimodal online instruction in supporting vocabulary acquisition among Chinese fifth-grade EFL learners. To address the limitations of traditional rote-based vocabulary teaching, the researcher implemented an eight-session extracurricular English course that integrated visual, auditory, and interactive modalities. Tasks included image matching, video clips, role-playing, digital games, and drawing-based activities.

A mixed-methods design was used to collect data through vocabulary pre- and post-tests, weekly formative homework assignments, student journals, classroom observations, and semi-structured interviews. The participants were 12 public school students aged 10–12 with varying levels of English proficiency. Results showed an average vocabulary score improvement of 11.58 percentage points, with the most notable gains among mid-level learners. Additionally, students displayed significantly higher engagement during game-based and image-supported tasks.

Qualitative data revealed that most students experienced positive emotions and greater confidence during multimodal lessons, particularly when learning was framed as interactive and playful. While some students initially struggled with technology, their motivation increased over time.

A majority of students (60%) exhibited positive emotional responses to the multimodal tasks, particularly interactive activities such as Kahoot, role-playing, and drawing. These tasks were found to be highly effective for students who demonstrated enthusiasm and confidence. For example, one student expressed, "This is better than school English class. I can speak more without worrying." The positive emotional reactions of these students were strongly correlated with higher participation levels and increased vocabulary retention.

In particular, students with positive emotional responses seemed to engage more deeply with tasks involving visual and kinesthetic elements. These activities not only helped students remember words more effectively but also gave them the confidence to apply the new vocabulary in practical, communicative contexts. This aligns with Mayer's (2009) dual-channel theory, which suggests that the integration of visual and auditory inputs enhances retention. Moreover, the use of game-based tasks allowed students to practice vocabulary in an enjoyable and risk-free environment, further reinforcing their learning. These findings underscore that emotionally engaged students performed better in vocabulary tasks, particularly those that required active involvement and interaction.

NEGATIVE EMOTIONAL RESPONSES AND LEARNING CHALLENGES

Around 15% of students displayed negative emotional responses, often due to anxiety related to speaking tasks or difficulties with the technology used in the lessons. One student shared, "I feel nervous to answer," reflecting a common challenge for these students. The negative emotional responses of these students were linked to decreased participation in tasks, especially those involving speaking and performance, which required more confidence and comfort with the language.

Interestingly, while these students made some progress in vocabulary acquisition, their improvement was much slower compared to their more emotionally engaged peers. These findings align with Zhao and Wang's (2020) study, which suggested that learners with higher anxiety levels or negative emotional responses may experience slower vocabulary acquisition despite exposure to multimodal input. For example, students who felt uneasy about technology or speaking tasks were more likely to avoid such tasks, thus limiting their opportunities for practice and language application.

To support these students in future cycles of instruction, strategies to reduce technical anxiety and boost student confidence are essential. Teachers could provide additional technical support, such as offering step-by-step guides or more practice sessions with technology before the main tasks. Additionally, role-playing and game-based activities should be designed with a gradual difficulty curve to give students the opportunity to build confidence and engage without feeling overwhelmed. Providing positive reinforcement and creating a safe environment for mistakes will help these students overcome their anxieties and improve their participation and retention. The results revealed the following emotional responses:

Positive Responses (60%): Many students enjoyed interactive activities such as games (Kahoot), role-playing, and drawing. These activities seemed to enhance their confidence, with one student stating, "This is better than school English class. I can speak more without worrying."

Neutral Responses (25%): Some students noted that the activities were "new" or "interesting," but did not express a strong preference for them. One student said, "I like the pictures, but I'm not sure I learn better."

Negative Responses (15%): A small group of students struggled with the technology or felt nervous about speaking. One student shared, "I feel nervous to answer."

Figure 3 below summarizes the distribution of emotional responses based on journal entries. This figure visually represents the breakdown of student emotional responses, providing further context to the subsequent analysis of how these emotions influenced student engagement and vocabulary acquisition.

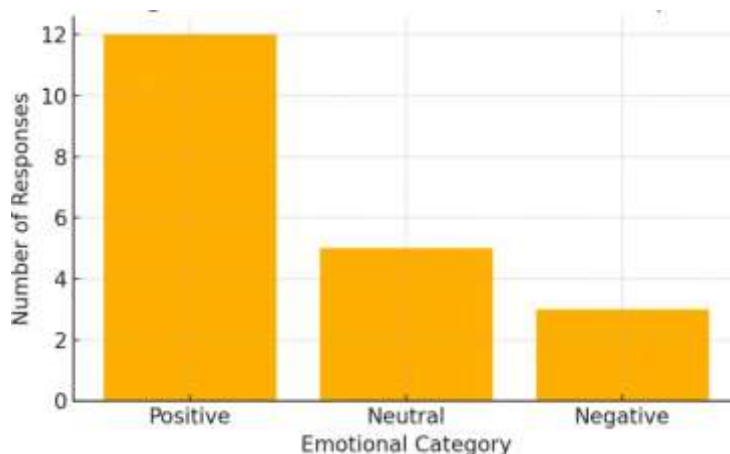


Figure 3: Categorization of Student Emotional Responses (n=20)

EMOTIONAL ENGAGEMENT AND INTRINSIC MOTIVATION

The relationship between emotional engagement and intrinsic motivation was particularly evident in this study. As Krashen (1982) suggests in his Affective Filter Hypothesis, learners' emotional states—such as anxiety, confidence, and enjoyment—directly influence their ability to acquire new language. When students felt more confident and emotionally engaged, they were more motivated to participate and practice the new vocabulary.

For students with positive emotional responses, their emotional investment in the learning process seemed to boost their intrinsic motivation. They were more likely to engage with the content for the sake of learning rather than simply completing the task. This is consistent with Plass et al. (2015), who emphasized that emotional engagement plays a crucial role in enhancing motivation and learning outcomes, especially in gamified and interactive environments. The increase in intrinsic motivation among these students contributed not only to better vocabulary retention but also to an overall more enjoyable and meaningful learning experience.

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

On the other hand, students with negative or neutral emotional responses were less likely to experience this boost in intrinsic motivation. Their participation tended to be more extrinsically driven (e.g., completing tasks for rewards or grades) rather than driven by a genuine interest in the learning material. As Krashen (1982) notes, this difference in emotional engagement can significantly impact how well students internalize new vocabulary. For these students, the lack of emotional involvement limited their overall motivation and slowed their progress in acquiring and applying new vocabulary.

IMPLICATIONS FOR TEACHING PRACTICE AND FUTURE CYCLES

The findings of this study underscore the importance of emotional engagement in enhancing students' vocabulary acquisition. To ensure that students with varying emotional responses can succeed, future teaching cycles should incorporate emotionally supportive strategies that address the emotional needs of students with both positive and negative emotional responses.

For students with positive emotional responses, it is essential to continue incorporating interactive, game-based tasks and kinesthetic activities, as these strategies foster deeper engagement and allow students to apply their learning in a more meaningful way. For students showing neutral or negative emotional responses, it is crucial to reduce emotional barriers by offering clear guidance, providing more practice with technology, and creating a safe space for students to make mistakes and build confidence. This could involve using scaffolding techniques, where students are given smaller, manageable tasks that gradually increase in difficulty.

INTERPRETATION AND LINK TO LITERATURE

The results of this study align with existing research on multimodal and gamified language teaching. The vocabulary gains observed here support the findings of Zhang and Zou (2021), who showed that multimodal input enhances word retention. The use of visual, auditory, and kinesthetic methods allowed students to engage with vocabulary in different ways, leading to stronger retention, especially for mid-level learners. Additionally, the increase in student engagement and participation aligns with studies by Zarzycka-Piskorz (2016) and Liu & Li (2022), who found that gamified learning increases motivation and fosters active participation. The role-playing and interactive tasks in this study played a key role in maintaining student interest and facilitating language use. The emotional responses of students also highlight the importance of affective factors in language learning. Plass et al. (2015) argue that reducing anxiety and fostering confidence are critical for language acquisition. In this study, the observed increase in student confidence, particularly among lower-level learners, suggests that multimodal learning can lower the affective filter, making language learning more accessible.

EDUCATIONAL IMPLICATIONS AND CONCLUSION

The integration of multimodal strategies in an online learning environment proved to be an effective method for vocabulary acquisition. The combination of visual aids, games, role-playing, and interactive tasks ensured that vocabulary was learned in meaningful, context-rich environments. This approach not only improved vocabulary retention but also enhanced student engagement and emotional investment in the learning process, which are critical for success in remote or hybrid learning contexts. Future teachers should consider incorporating multimodal approaches into their vocabulary instruction, especially for younger learners or those with varying proficiency levels. Tasks that blend visual, auditory, and kinesthetic elements are particularly beneficial in supporting diverse learning styles and promoting long-term retention.

REFLECTION

This action research project provided a valuable opportunity to reflect on my roles as an English teacher, a classroom researcher, and an instructional designer. Throughout the eight-session intervention, I encountered both expected and unexpected challenges and discovered practical insights that will influence my future teaching and research.

REFLECTION AS A TEACHER

One of the most profound realizations I had as a teacher was the importance of emotional engagement in language learning. In traditional EFL classes, vocabulary is often taught through repetition, memorization, and testing. While efficient, these methods rarely spark enthusiasm. In contrast, when I introduced multimodal tasks such as role-playing, image-based guessing games, and interactive tools like Kahoot, I noticed a visible shift in student behavior.

Students who were usually passive became more vocal. Some even initiated language use beyond my expectations. For instance, one low-level student volunteered to describe his drawing using multiple new words. These moments reminded me that confidence often grows through play, not pressure.

However, I also faced difficulties. Some students struggled with the technology, and initial class time was consumed by teaching them how to navigate platforms. This taught me to include "digital readiness" as a teaching priority in online environments.

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

REFLECTION AS A RESEARCHER

From a research perspective, I gained a deeper understanding of how to work with mixed-methods data. Initially, I thought the vocabulary test scores would be the clearest indicator of learning success. But I soon realized that qualitative data—student journals, interview quotes, observation logs—provided far richer insight.

For example, two students showed only slight improvement in scores but reported higher motivation, stronger self-confidence, and enjoyment of the learning process. Had I relied solely on test data, I would have underestimated the intervention's true impact.

I also learned the importance of consistency in formative assessments. Because my quizzes were completed at home, some students received help or referred to notes. This reminded me to be cautious in interpreting data collected under variable conditions.

REFLECTION AS AN INSTRUCTIONAL DESIGNER

Designing each lesson was perhaps the most intellectually demanding yet rewarding part of the research. My goal was to balance engagement with instructional goals—ensuring that fun activities were not just entertaining, but purpose-driven.

Over time, I developed a "task chain" approach: beginning each lesson with input (video, visuals), followed by structured practice (games or matching), and ending with output (role-play, speaking, or creative tasks). This progression helped students internalize vocabulary through repeated, contextualized exposure.

I also learned to adapt. In one lesson, my planned storytelling activity took too long, so I replaced it with a drawing-and-describe task that better fit student energy levels. Such in-the-moment adjustments were key to maintaining momentum in online teaching.

This action research not only improved my students' vocabulary learning but also transformed the way I view teaching. It made me more observant, more reflective, and more intentional. Moving forward, I am committed to designing English learning experiences that are not just educational—but also joyful, inclusive, and emotionally safe.

CONCLUSION AND RECOMMENDATIONS

This action research investigated the effects of multimodal online instruction on vocabulary acquisition and learner engagement among Chinese fifth-grade EFL students. Through an eight-session extracurricular course integrating visual, auditory, and interactive elements, the study found that students not only improved their vocabulary scores but also demonstrated higher motivation and emotional involvement.

The results indicated that multimodal input—such as games, drawings, videos, and role-plays—can significantly support vocabulary learning, particularly when instruction is context-rich and emotionally engaging. Even low-proficiency students showed increased willingness to participate and express themselves. These findings reinforce the pedagogical value of learner-centered, multimodal approaches in digital language teaching environments.

RECOMMENDATIONS

Based on the findings and classroom experience, the following recommendations are proposed for teachers and future researchers:

1. Prepare students for digital learning: Before introducing multimodal tasks, provide basic training on how to use interactive tools like Kahoot or digital whiteboards to reduce anxiety and save time during lessons.
2. Design purposeful, progressive tasks: Ensure that each activity builds on the last, forming a logical task chain from input to practice to output. Activities should not only be fun but serve clear language goals.
3. Balance assessment methods: While homework can be useful, in-class formative assessments are more reliable and better reflect student understanding without external assistance.
4. Explore peer collaboration: Future implementations may consider including small-group projects or peer feedback to enhance communicative use of vocabulary.
5. Extend the research scope: Future studies could compare multimodal instruction across different age groups or school types, or investigate long-term retention effects.

REFERENCES

- 1) Chen, X., & Huang, Y. (2021). The impact of visual aids on students' vocabulary learning in EFL contexts. *Journal of Language Teaching and Learning*, 45(3), 233-245. <https://doi.org/10.1016/j.jlt.2021.04.005>
- 2) Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining "gamification". *Proceedings of the 2011 annual conference on Human factors in computing systems* (pp. 2425–2428). Association for Computing Machinery. <https://doi.org/10.1145/1978942.1979299>
- 3) Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- 4) Liu, M., & Li, L. (2022). Gamification in language learning: A study on the role of competitive games in EFL classrooms. *Language Learning and Technology*, 26(1), 72-87. <https://doi.org/10.1016/j.llet.2022.02.003>

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

- 5) Mayer, R. E. (2009). Multimedia learning (2nd ed.). Cambridge University Press. Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. Educational Psychologist, 50(4), 258-283. <https://doi.org/10.1080/00461520.2015.1122533>
- 6) Zhang, M., & Zou, L. (2021). Multimodal learning and its effect on vocabulary acquisition in young learners: Evidence from a Chinese primary school study. TESOL Quarterly, 55(2), 413-437. <https://doi.org/10.1002/tesq.457>
- 7) Zhao, Y., & Wang, T. (2020). The role of contextualized multimodal input in vocabulary acquisition. Language Teaching Research, 24(1), 89-105. <https://doi.org/10.1177/1362168819875334>

APPENDIX

Formative Test: Unit 1 - My School

Name: _____ Class: _____ Date: _____

Part 1: Listening (听力测试) - 10 points

A. Listen and Circle (听音, 圈出正确的单词) (5 points)

1. book / bag
2. chair / table
3. crayon / ruler
4. desk / window
5. eraser / sharpener

B. Listen and Match (听音, 连线) (5 points)

Audio reads sentence like: "It's a pen." Students match the sentence to the correct picture.



Part 2: Vocabulary (词汇测试) - 10 points

A. Look and Write (看图写单词) (5 points)

Students see images of **book, chair, crayon, desk, eraser** and write the correct word below each picture.

B. Unscramble the Words (单词重组) (5 points)

Rearrange the letters to form the correct words:

1. okob → _____
2. iarch → _____
3. anrcoy → _____
4. ksde → _____
5. srreca → _____

Part 3: Sentence Patterns (句型测试) - 10 points

A. Fill in the Blanks (填空) (5 points)

Complete the sentences using the correct words:

1. Q: What is it?
A: It's a _____. (pen)
2. Q: What is it?
A: It's a _____. (desk)
3. Q: _____ is it?
A: It's a crayon.
4. Q: What _____ it?
A: It's a book.
5. Q: What is it?
A: It's a _____. (eraser)

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

B. Matching (连线) (5 points)

Match the questions with the correct answers.

1. What is it? → **A. It's a pencil.**
2. What is it? → **B. It's a chair.**
3. What is it? → **C. It's an eraser.**
4. What is it? → **D. It's a desk.**
5. What is it? → **E. It's a book.**

Part 4: Writing (书写练习) - 10 points

Look at the picture and write a full sentence using "What is it? - It's a _____."

Example:



Q: What is it?

A: It's a pen.

Formative Test: Unit 2 - My Toys

Name: _____ Class: _____ Date: _____

Part 1: Listening (听力测试) - 10 points

A. Listen and Circle (听音, 圈出正确的单词) (5 points) ball / bear

kite / car robot / puzzle plane / doll

car / teddy bear

B. Listen and Match (听音, 连线) (5 points)

音频内容如: "It's a kite." 学生将句子与对应图片连线。

Part 2: Vocabulary (词汇测试) - 10 points

A. Look and Write (看图写单词) (5 points)

展示以下玩具图片, 学生写出对应单词: ball, doll, robot, kite, car

B. Unscramble the Words (单词重组) (5 points) lalb → _____

drlo → _____

tkei → _____

obrot → _____

rca → _____

Part 3: Sentence Patterns (句型测试) - 10 points

A. Fill in the Blanks (填空) (5 points) Q: What do you want?

A: I want a _____. (robot) Q: What do you want?

A: I want a _____. (ball)

Q: What do you _____?

A: I want a doll.

Q: What do you want? A: I _____ a car. Q: What do you want?

A: I want a _____. (kite)

B. Matching (连线) (5 points)

将问题与答案连线:

What do you want? → A. I want a teddy bear. What do you want? → B. I want a kite.

What do you want? → C. I want a puzzle. What do you want? → D. I want a robot. What do you want? → E. I want a car.

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

Part 4: Writing (书写练习) - 10 points

看图写完整句子, 使用 "What do you want? — I want a _____."

Example: (图片: A teddy bear) Q: What do you want?

A: I want a teddy bear.

Formative Test:

Unit 3 - My Family

Name: _____ Class: _____ Date: _____

Part 1: Listening (听力测试) - 10 points

A. Listen and Circle (听音, 圈出正确的单词) (5 points)

mother / father sister / brother grandma / grandpa aunt / uncle cousin / brother

B. Listen and Match (听音, 连线) (5 points)

Audio example: "This is my sister." Students match the sentence to the correct picture.

Part 2: Vocabulary (词汇测试) - 10 points

A. Look and Write (看图写单词) (5 points) Words: mother, father, sister, brother, grandpa

B. Unscramble the Words (单词重组) (5 points) omther → _

arthfe → _____

tsesir → _____

rgopand → _____

brthor → _____

Part 3: Sentence Patterns (句型测试) - 10 points

A. Fill in the Blanks (填空) (5 points) Q: Who's this?

A: It's my _____. (mother) Q: Who's this?

A: It's my _____. (grandfather) Q: _____ this?

A: It's my sister.

Q: Who's this?

A: It's my _____. (father) Q: Who's this?

A: It's my _____. (brother)

B. Matching (连线) (5 points) Who's this? → A. It's my grandma.

Who's this? → B. It's my aunt. Who's this? → C. It's my cousin. Who's this? → D. It's my sister. Who's this? → E. It's my grandpa.

Part 4: Writing (书写练习) - 10 points

Look at the picture and write a sentence using: "Who's this? — It's my _____."

Formative Test:

Unit 4 - My Body

Name: _____ Class: _____ Date: _____

Part 1: Listening - 10 points

A. Listen and Circle (5 points) head / hand

nose / eyes leg / ear foot / finger arm / knee

B. Listen and Match (5 points)

Audio example: "This is my nose." Students match with image.

Part 2: Vocabulary - 10 points

A. Look and Write (5 points) Words: head, nose, arm, eyes, foot

B. Unscramble the Words (5 points)

edah → _____

esno → _____

mra → _____

seey → _____

otfo → _____

Part 3: Sentence Patterns - 10 points

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

A. Fill in the Blanks (5 points) I have two _____.

B. (hands) She has one _____.

C. (nose) He has two _____.

D. (eyes) I _____ two legs.

E. (have) He _____ one mouth. (has) Matching (5 points)

What do you have? → A. I have two ears. What does she have? → B. She has two arms. What does he have? → C. He has one nose.

What do they have? → D. They have two legs. What do you have? → E. I have a head.

Part 4: Writing - 10 points

Use the structure: "I have two _____." / "He has one _____."

Formative Test:

Unit 5 - Stories

Name: _____ Class: _____ Date: _____

Part 1: Listening - 10 points

A. Listen and Circle (5 points) bird / crown king / queen prince / princess happy / sad story / friends

B. Listen and Match (5 points)

Example: "She is a princess." Match with correct picture.

Part 2: Vocabulary - 10 points

A. Look and Write (5 points) Words: king, queen, bird, crown, happy

B. Unscramble the Words (5 points)

gkin → _____

nequee → _____

drbi → _____

owncrc → _____

yppah → _____

Part 3: Sentence Patterns - 10 points

A. Fill in the Blanks (5 points) She is a _____. (princess)

He is a _____. (king)

The _____ is happy. (queen) I see a _____. (bird)

The prince is _____. (sad)

B. Matching (5 points)

Who is she? → A. She is a princess. Who is he? → B. He is a prince.

Is she happy? → C. Yes, she is.

Is the king sad? → D. No, he is happy. What is it? → E. It's a crown.

Part 4: Writing - 10 points

Use the sentence: "He is a _____." or "She is a _____." to describe characters.

Formative Test:

Unit 6 - I Like Food

Name: _____ Class: _____ Date: _____

Part 1: Listening - 10 points

A. Listen and Circle (5 points) bread / rice milk / orange juice chicken / noodles cookies / bananas water / soup

B. Listen and Match (5 points) Example: "I like milk." Match with image.

Part 2: Vocabulary - 10 points

A. Look and Write (5 points) Words: noodles, milk, bread, rice, bananas

B. Unscramble the Words (5 points)

ikml → _____

drebad → _____

ceir → _____

donloes → _____

ssankaba → _____

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

Part 3: Sentence Patterns - 10 points

A. Fill in the Blanks (5 points) Q: What do you like?

A: I like _____. (bread) Q: Do you like rice?

A: _____. (Yes, I do.) Q: Do you like milk?

A: _____. (No, I don't.) Q: What _____ you like?

A: I like noodles.

A: I like _____. (chicken)

B. Matching (5 points)

Do you like bananas? → A. Yes, I do. Do you like cookies? → B. No, I don't. What do you like? → C. I like milk.

Do you like water? → D. Yes, I do.

What do you like? → E. I like rice.

Part 4: Writing - 10 points

Write about what you like.

Example: "I like cookies." / "I like orange juice."

Formative Test:

Unit 7 - Clothes

Name: _____ Class: _____ Date: _____

Part 1: Listening - 10 points

A. Listen and Circle (5 points) coat / dressshoes / socks shirt / shorts pants / skirt hat / cap

B. Listen and Match (5 points)

Audio reads: "These are my shoes." Match with picture.

Part 2: Vocabulary - 10 points

A. Look and Write (5 points) Words: coat, dress, pants, shirt, hat

B. Unscramble the Words (5 points)

tcoaa → _____

sssok → _____

sirht → _____

sartn → _____

kehatr → _____

Part 3: Sentence Patterns - 10 points

A. Fill in the Blanks (5 points) These are my _____.

B. (pants) This is my _____. (hat)

Are these your shoes? Yes, _____.

Is this your dress? No, _____.

I like your _____. (shirt)

C. Matching (5 points)

What is this? → A. It's my coat.

What are these? → B. They are my socks. Is this your dress? → C. Yes, it is.

Are these your pants? → D. No, they aren't. I like your skirt. → E. Thank you!

Part 4: Writing - 10 points

Describe your clothes: "This is my _____." / "These are my _____."

Formative Test:

Unit 8 - Animals

Name: _____ Class: _____ Date: _____

Part 1: Listening - 10 points

A. Listen and Circle (5 points) lion / monkey zebra / giraffe crocodile / hippo drink / eat run / walk

B. Listen and Match (5 points)

Example: "The lion can run." Match with picture.

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

Part 2: Vocabulary - 10 points

A. Look and Write (5 points) Words: lion, giraffe, monkey, zebra, hippo

B. Unscramble the Words (5 points)

onil → _____

pphio → _____

knyemo → _____

afgierf → _____

zberaa → _____

Part 3: Sentence Patterns - 10 points

A. Fill in the Blanks (5 points) The lion can _____. (run) The hippo can. (walk) Can the giraffe run? Yes, _____
What animal is it? It's a _____. (zebra) The monkey can _____. (eat)

B. Matching (5 points)

What is it? → A. It's a crocodile. Can the lion run? → B. Yes, it can.

What can the hippo do? → C. It can walk. What is it doing? → D. It's drinking. _____

Can the giraffe swim? → E. No, it can't.

Part 4: Writing - 10 points Write about an animal:

Example: "The zebra can run." / "The crocodile can eat."

Pre-Test:

Name: _____ Class: _____ Date: _____

Part 1: Vocabulary – 10 points

A. Choose the correct word for the picture. (看图，选单词)

老师出示实物图或 PPT，学生在下面选出正确单词。（此处可配图）

A. book B. pen C. robot

A. kite B. doll C. mother

A. milk B. rice C. skirt

A. lion B. giraffe C. zebra

A. socks B. shoes C. shirt

B. Circle the word you know. (圈出你认识的单词)

6. coat / queen

7. banana / uncle

8. desk / hippo

9. plane / nose

10. teddy bear / pencil

Part 2: Sentence Patterns – 10 points

A. Match the questions with answers. (连线) What is it? → A. It's a kite.

Who's this? → B. I like rice.

What do you want? → C. It's my sister. Do you like milk? → D. I want a robot. What do you like? → E. Yes, I do.

Part 3: Writing – 10 points Look and Write (看图写句子)

老师出示图片，学生模仿例句完成句子（也可口头辅助）

(图片: apple)

→ I like _____. (图片: mother)

→ This is my _____. (图片: shoes)

→ These are my _____. (图片: lion)

→ The lion can _____. (图片: pencil)

→ What is it? It's a _____.

Post-Test: Name: _____ Class: _____ Date: _____

Part 1: Listening (听力测试) – 10 points

A. Listen and Circle (5 points) 听录音，圈出你听到的单词: ruler / rubber teddy bear / puzzle grandma / grandpa rice / bread giraffe / hippo

B. Listen and Choose the Right Answer (5 points) 听句子并选择最合适的图片或句子回答。如: "Who's this?" → A. It's my

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

sister. B. I like juice. C. I have socks.

Part 2: Vocabulary (词汇测试) – 10 points

A. Choose the Right Word (5 points) 选择与图或提示相符的单词:

_____ (used to eat) → A. pen B. spoon C. shoe

_____ (flies in the sky) → A. ball B. kite C. robot

_____ (part of the body) → A. hand B. hat C. hammer

_____ (animal with long neck) → A. monkey B. lion C. giraffe

_____ (food) → A. milk B. queen C. ruler

B. Categorize the Words (5 points) 将以下词汇归类: coat, banana, pencil, hippo, sister 分类:

Clothes: _____ Food: _____ School: _____ Animals: _____ Family: _____

Part 3: Sentence Patterns (句型测试) – 10 points

A. Complete the Dialogues (5 points) 补全对话句子:

A: What is it?

B: It's a _____.

A: Do you like noodles? B: _____.

A: Who's this?

B: _____.

A: What do you want? B: _____.

A: Can the monkey climb? B: _____.

B. Reorder the Words (5 points)

将乱序的词语重组为完整句子:

I / ball / want / a → _____

my / This / brother / is → _____

have / I / two / eyes → _____

queen / is / She / a → _____

can / The / lion / run → _____

Part 4: Writing (书写练习) – 10 points

根据提示词写句子。

提示词: robot, dress, grandma, bananas, zebra I want a _____.

This is my _____.

I like _____.

The _____ is eating.

She has a _____.

Parental/Guardian Consent Form

家长/监护人同意书

Dear Parent/Guardian,

尊敬的家长/监护人:

I am Heng Zhang, an English as a Foreign Language (EFL) teacher. I am conducting a research study as part of my coursework to complete my master's degree program in De La Salle University. This research aims to explore how "multimodal online teaching" (using videos, images, audio, and interactive activities) helps Chinese fifth-grade EFL learners acquire vocabulary and understand their experiences with this teaching approach. The findings will help improve online teaching methods and support students in learning English more effectively.

您好! 我是张恒, 一名英语作为外语 (EFL) 教师。我正在进行一项研究, 以完

成我在德拉萨大学硕士学位课程的学习。本研究旨在探讨"多模态在线教学"(使用视频、图片、音频和互动活动) 如何帮助中国五年级 EFL 学习者习得词汇, 并了解他们对这种教学方式的体验。研究结果将有助于改进在线教学方法, 从而更有效地支持学生学习英语。

Study Procedures

1. Students will participate in multimodal online teaching sessions twice a week for four (4) weeks. These sessions will include

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

videos, images, audio, and interactive activities to teach new vocabulary.

2. Students will take a pre-test before the study begins and a post-test after the study to measure their vocabulary acquisition.
3. During the teaching sessions, students will complete formative assessments to monitor their learning progress and provide timely feedback.
4. Four (4) students will be selected for a short interview to share their experiences with multimodal online teaching.
5. I will also use an observation checklist during the sessions to note students' engagement and participation.

研究流程

1. 学生将在四（4）周内每周参加两次多模态在线教学课程。这些课程将通过视频、图片、音频和互动活动来教授新词汇。
2. 学生将在研究开始前进行前测（pre-test），在研究结束后进行后测（post-test），以测量他们的词汇习得情况。
3. 在教学过程中，学生将完成形成性评估，以监测学习进度并提供及时反馈。
4. 将随机选取四（4）名学生进行简短访谈，以分享他们对多模态在线教学的体验。
5. 在教学过程中，我还将使用观察清单来记录学生的参与度和互动情况。

Duration

The study will be conducted for approximately four (4) weeks for the online teaching sessions. The pre-test will be administered before the study begins, and the post-test will be conducted two weeks after the last session. Formative assessments will be administered regularly during the teaching sessions, and interviews will take place during the final week of the study.

研究时长

本研究的在线教学课程将持续约四（4）周。在研究开始前将进行前测（pre-test），而后测（post-test）将在最后一节课程结束后两周进行。形成性评估将在教学过程中定期进行，而访谈将在研究的最后一周进行。

Confidentiality

No personal information will be shared. Students' names will not appear in any report. Pseudonyms will be used for the participants, and all data will be kept confidential. The research is only for educational purposes.

保密性

所有个人信息均不会被公开。学生的姓名不会出现在任何报告中，研究过程中将使用化名来代替真实姓名，所有数据都将严格保密。本研究仅用于教育目的。

Voluntary Participation

The participation of your child in this research study is voluntary. If you do not give consent, your child will not be included in the study. If you decide to let your child participate, you will be asked to sign this consent form. You or your child can stop participation at any time without any problem.

自愿参与

您孩子参与本研究是自愿的。如果您不同意，您的孩子将不会被纳入研究。如果您同意让您的孩子参加，您需要签署本同意书。您或您的孩子可以在任何时候退出研究，而不会有任何影响。

How Data Will Be Used

1. Pre-test, formative assessment, and post-test scores will be analyzed using descriptive statistics to understand how multimodal online teaching facilitates vocabulary acquisition.
2. Interview transcripts will be analyzed thematically to identify key themes related to students' experiences with multimodal online teaching. 数据使用方式

1. 前测、形成性评估和后测的分数将通过描述性统计分析，以了解多模态在线教学如何促进词汇习得。
2. 访谈记录将进行主题分析，以识别与学生多模态在线教学体验相关的关键主题。

Student Information | 学生信息

- **Student Name (学生姓名):** _____
- **Date of Birth (出生日期):** _____

An Action Research on Multimodal Online Teaching and Vocabulary Acquisition Among Chinese Fifth-Grade EFL Learners

• **Class/Grade (班级/年级):** _____

Guardian Information | 监护人信息

Guardian Name (监护人姓名): _____

Relationship to Student (与学生关系): _____

Phone Number (联系电话): _____

Consent Statement | 同意声明

I, [Guardian's Name], as the legal guardian of [Student's Name], give my full consent for my child to participate in this research. I understand the nature and purpose of this activity and acknowledge that all reasonable precautions will be taken to ensure the safety of my child.

本人, [监护人姓名], 作为[学生姓名]的法定监护人, 同意我的孩子参加这项研究。我理解该活动的性质和目的, 并承认主办方会采取所有合理的预防措施, 以确保我的孩子的安全。

Example of a Kahoot! Quiz Question Used in Class



There is an Open Access article, distributed under the term of the Creative Commons Attribution – Non Commercial 4.0 International (CC BY-NC 4.0) (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits remixing, adapting and building upon the work for non-commercial use, provided the original work is properly cited.