

Designing a Theory-Driven Multimodal Vocabulary Learning Module for Chinese EFL Learners

Yaqiong Meng

Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor 43600, Malaysia

Supyan Hussin

Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor 43600, Malaysia

Harwati Hashim

Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor 43600, Malaysia

Abstract—This study presents the design of a theory-driven multimodal vocabulary learning module developed for first-year non-English majors at a Chinese application-oriented university. Drawing upon Multimodal Literacy Framework, Discourse-Based Instruction, and the Gradual Release of Responsibility (GRR) model, the module aims to enhance learners' productive vocabulary knowledge through structured, scaffolded instruction that integrates visual, auditory, kinesthetic, and textual modalities. The development process followed the Sidek Module Development Model, encompassing comprehensive steps from needs analysis to content selection, strategy integration, and instructional sequencing. Each activity within the module was explicitly aligned with the GRR instructional stages ("I do, We do, You do") to promote gradual learner autonomy in vocabulary output. The resulting module features a combination of vocabulary presentation, collaborative tasks, discourse-based reading, and multimodal production assignments such as storytelling and digital design. The study details the design rationale, theoretical underpinnings, and pedagogical structure of the module, providing a replicable framework for EFL educators seeking to foster productive vocabulary development in low-proficiency learners. While no empirical data are presented in this paper, the module is positioned for future classroom validation and practical adaptation across applied university settings.

Index Terms—multimodal learning, vocabulary instructional design, Gradual Release of Responsibility (GRR), Sidek Module Development Model, module design

I. INTRODUCTION

Over the past decade, higher education worldwide has increasingly emphasized employability to meet the evolving demands of the labor market (Scandurra et al., 2024). In China, this shift is reflected in the strategic promotion of application-oriented universities—institutions designed not only to transmit academic knowledge but also to serve regional economic development and industrial upgrading through practical skills training, innovation-driven pedagogy, and close university–industry collaboration (Zhao et al., 2024). In line with this mission to integrate education with industry and prepare graduates for workplace demands, college English instruction is undergoing a parallel transformation (Jiang, 2020; Liang, 2021; Meng et al., 2023). Rather than serving merely as a vehicle for linguistic knowledge acquisition and examination preparation, English courses are increasingly positioned as practical, skills-oriented components of the curriculum that enable students to communicate effectively in authentic academic and professional contexts (Fu, 2024). However, in many application-oriented universities, traditional practices—such as the grammar-translation method and rote vocabulary memorization—remain prevalent, limiting learners' capacity for real-time expression and meaningful engagement (Xu & Zhang, 2021).

Among the various components of English instruction, vocabulary learning remains both fundamental and underdeveloped. Most existing curricula focus heavily on receptive vocabulary—the ability to recognize and understand words—while placing far less emphasis on productive vocabulary (Pan, 2023), which involves the ability to accurately and appropriately use words in speaking and writing (Jia, 2024). Without sufficient productive vocabulary, learners struggle to express ideas, sustain interactions, and meet the communicative demands of academic or professional tasks (Schmitt et al., 2020; Webb & Nation, 2017).

To address this issue, multimodal vocabulary instruction has gained increasing scholarly attention in recent years. Drawing on the Cognitive Theory of Multimedia Learning (Mayer, 2005) and the framework of Multimodal Literacy (Jewitt, 2008; Kress & Van Leeuwen, 2001), this approach integrates visual, auditory, kinesthetic, and textual channels to enhance cognitive processing, memory retention, and learning depth. Empirical studies suggest that multimodal strategies can outperform unimodal methods in supporting vocabulary acquisition, especially among low-proficiency or

demotivated learners (Aedo & Millafilo, 2022; Bansong et al., 2023; Bai, 2023). Additionally, discourse-based instruction has been found to deepen word learning by embedding vocabulary in meaningful, real-world contexts (Gee, 2008; Hyland, 2014), bridging the gap between isolated vocabulary exposure and contextualized usage.

Yet, while both multimodal and discourse-based instruction offer promising directions, their classroom implementations often lack structure and pedagogical coherence. Many current practices are fragmented and short-term, with limited theoretical integration and instructional scaffolding (Argyriou & Tapsis, 2025; Huang, 2025; Muhammad Nur Iman & Nurasia Natsir, 2025). In particular, few designs systematically support learners' transition from guided input to independent output—a core process in vocabulary development.

To respond to these challenges, this study introduces a theory-driven vocabulary learning module that integrates Multimodal Literacy Framework, Discourse-Based Instruction, and the Gradual Release of Responsibility (GRR) model. Multimodal Literacy Framework informs the design by incorporating visual, auditory, kinesthetic, and textual channels to enhance learners' cognitive processing, memory retention, and engagement (Argyriou & Tapsis, 2025; Huang, 2025; Jewitt, 2008; Kress & Van Leeuwen, 2001). Discourse-Based Instruction situates vocabulary learning within meaningful, real-world contexts, enabling students to encounter and use target words in authentic communicative tasks (Gee, 2008; Hyland, 2014). GRR provides a structured instructional architecture—"I do, We do, You do"—that gradually shifts responsibility from teacher to learner, ensuring a smooth transition from guided practice to independent production (Fisher & Frey, 2021). Combined, these frameworks create a module that is multimodal and context-rich, while also being scaffolded and learner-centered, enabling students to build productive vocabulary knowledge with both theoretical grounding and practical relevance.

Methodologically, the module was developed using the Sidek Module Development Model (Madihie & Noah, 2013; Zainudin et al., 2021), a structured 13-step framework that ensures theoretical alignment, instructional feasibility, and design validity. The resulting module features sequenced vocabulary tasks, real-world discourse integration, and multimodal output activities supported by teacher modeling and feedback.

Instead of presenting empirical outcomes, this paper focuses on detailing the theoretical foundations, pedagogical rationale, and instructional structure of the module. It aims to provide a replicable and adaptable framework for EFL educators seeking to develop productive vocabulary knowledge among learners in applied university contexts.

II. LITERATURE REVIEW

This literature review provides the conceptual and empirical foundation for the theory-driven multimodal vocabulary learning module proposed in this study. The review is organized into four thematic areas: (1) productive vocabulary knowledge in EFL contexts, (2) multimodal vocabulary instruction, (3) discourse-based vocabulary instruction, and (4) instructional design models. Each section discusses key developments in the respective field and identifies research gaps relevant to the context of applied universities in China.

A. Productive Vocabulary Knowledge in EFL Contexts

Productive vocabulary knowledge refers to learners' ability to retrieve and use lexical items appropriately in speaking and writing. It requires not only knowledge of word form and meaning, but also accurate use in context through grammatical, collocational, and register awareness (Nation, 2006; Schmitt et al., 2020). Compared to receptive vocabulary, productive vocabulary is cognitively more demanding as it involves lexical recall, contextual adaptation, and real-time generation (Laufer, 2005; Laufer & Nation, 1999).

In the past decade, researchers have increasingly highlighted the gap between receptive and productive vocabulary development, especially among EFL learners. Empirical studies show that productive vocabulary is best developed through output-oriented and context-rich tasks. For example, Teng and Xu (2022) demonstrated that translation, retelling, and generative writing tasks significantly improved learners' ability to produce target words. Similarly, Jia (2024) found that productive use of vocabulary in speaking and writing tasks correlates strongly with communicative performance, more so than recognition-based knowledge. Repetition, contextualization, and output opportunities have emerged as critical variables in recent findings (Pan, 2023; Teng & Xu, 2022).

Despite these advances, few studies have systematically applied these strategies in the setting of Chinese applied universities. These institutions often face unique challenges such as limited instructional time, low learner motivation, and lack of authentic task contexts, which hinder effective vocabulary development. Therefore, there is a pressing need to develop structured and context-sensitive interventions that support productive vocabulary growth in these settings.

B. Multimodal Vocabulary Instruction: From Theory to Practice

Building on the need for more effective vocabulary development methods, multimodal instruction has emerged as a promising approach that supports learners' lexical development through multiple sensory and symbolic channels. It addresses key cognitive challenges in vocabulary learning by integrating input across modalities and enhancing retention, engagement, and depth of processing.

Multimodal instruction is grounded in Mayer's (2005) cognitive theory of multimedia learning and Kress and van Leeuwen's (2001) theory of multimodal literacy. It emphasizes the pedagogical value of combining visual, auditory, kinesthetic, and textual input to facilitate meaning construction. Early research has shown that multimodal learning

strategies, which involve multiple sensory systems simultaneously, can significantly enhance vocabulary retention and comprehension. For instance, Al-Jarf (2023) demonstrated that integrating visual stimuli such as images and animations into vocabulary instruction can significantly improve students' ability to remember and use new words.

Recent developments have focused on learner-generated multimodal tasks, made possible by digital tools like Quizlet, Canva, and video editing platforms. Teng (2023) showed that such tools improve learners' attention, emotional engagement, and processing depth. Bansong et al. (2023) confirmed that creating visual stories and multimedia compositions helps learners engage more deeply with vocabulary, making it more meaningful and memorable.

Yet, empirical syntheses reveal that classroom implementations of multimodal or discourse-based instruction often lack structure and pedagogical coherence. Argyriou and Tapsis (2025) found that multimodal research remains fragmented across disciplines and is typically delivered as short-term, isolated interventions without sustained scaffolding. Iman and Natsir (2025) reported that even when translanguaging—an inherently multimodal practice—is introduced in digital environments, teachers struggle to integrate pedagogical aims with technological affordances, resulting in sporadic and disjointed activities. Huang (2025), reflecting on Chinese EFL contexts, observed that the newly adopted “Smart Classrooms” frequently reduce multimodality to a display of digital devices rather than a coherent instructional sequence, leading to what she terms “formalism without pedagogical substance.” Collectively, these studies underscore the urgent need for theoretically integrated, replicable frameworks that embed multimodal techniques within sustained, scaffolded module designs aligned with learners' cognitive and contextual needs.

C. Discourse-Based Vocabulary Instruction: Bridging Context and Use

While multimodal instruction can enhance cognitive input and learner engagement by integrating multiple sensory channels—visual, auditory, kinesthetic, and textual—multimodality alone is not sufficient to ensure the effective use of vocabulary in authentic communication. Discourse-based vocabulary instruction complements this approach by embedding target vocabulary within real-world genres and communicative contexts, thereby increasing the contextualization and functional relevance of learning. Caro and Mendinueta (2017) pointed out that lexical competence encompasses not only knowledge of word meanings but also mastery of collocations, grammatical patterns, and discourse functions. Since the mid-2010s, research has shown that embedding vocabulary learning in authentic tasks—such as narrative writing, workplace emails, and academic presentations—can significantly enhance learners' vocabulary mastery and usage (Webb & Nation, 2017). Building on this, Wakasa (2024) demonstrated that discourse analysis not only improves ESL learners' reading comprehension but also, by increasing contextualization and interactivity, enables teachers to better tailor instruction to learners' communicative needs and to promote a deeper understanding of vocabulary use. Similarly, Qiann et al. (2025) found that integrating instructional discourse analysis with communicative language teaching not only improves students' pragmatic competence but also transforms classroom interaction patterns, making learning more engaging and inclusive; students participating in such activities exhibited greater fluency and confidence in authentic communication.

The integration of multimodality and discourse analysis offers a complementary and synergistic pathway for vocabulary instruction. On the one hand, multimodal tasks stimulate learners' creative expression and active engagement; on the other, discourse-driven activities ensure the contextual authenticity and functional appropriateness of vocabulary use. Research by Teng (2023) shows that when learners create multimodal artifacts such as infographics, role-plays, or video blog scripts, they not only enhance their vocabulary and discourse competence but also strengthen creativity and motivation. Such integrated tasks are particularly well suited to the instructional priorities of application-oriented universities, where professional relevance and communicative competence are emphasized. By combining the two approaches, teachers can create a rich, interactive, and application-oriented learning environment that supports the simultaneous development of vocabulary acquisition and overall communicative competence.

D. Instructional Design Models: The Role of the GRR and Sidek Module Development Model

The effective integration of multimodal and discourse-based vocabulary instruction requires a structured instructional design framework. While the ADDIE model (Branch, 2009) is widely used, it has been criticized for its generality and limited consideration of learner diversity and early-stage validation (Simmons, 2012).

To address these limitations, this study adopts two complementary frameworks: the Gradual Release of Responsibility (GRR) model and the Sidek Module Development Model (SMDM). The GRR model (Fisher & Frey, 2021) provides a scaffolded instructional delivery structure that guides the implementation of each activity. Organized into three sequential stages—“I do,” “We do,” and “You do”—the model supports students' gradual transition from teacher modeling to collaborative practice and, ultimately, to independent vocabulary use. GRR is particularly well-suited for structured output-oriented vocabulary learning in EFL contexts, as it ensures learners receive sufficient support before being asked to perform autonomously.

At a broader level, the Sidek Module Development Model (Madihie & Noah, 2013) offers a systematic, theory-informed framework for the development of the vocabulary learning module itself. Originally designed for counseling interventions, SMDM has since been applied to educational settings, including language learning (Zainudin et al., 2021). The model emphasizes needs analysis, cultural contextualization, and empirical validation. SMDM consists of two main phases. The first, module development, includes nine steps: (1) aim setting, (2) theory and learner identification, (3) needs analysis, (4) goal setting, (5) content selection, (6) strategy selection, (7) logistics planning, (8) media and

technology integration, and (9) component integration. The second phase, validation, involves expert review and pilot testing to ensure content validity and instructional effectiveness.

Together, GRR and SMDM serve complementary roles: GRR structures the internal instructional flow of each activity, while SMDM guides the overarching development and evaluation of the module. This dual-framework approach ensures both pedagogical coherence and developmental rigor in the implementation of the multimodal vocabulary learning module.

In summary, research in vocabulary instruction has shifted from rote memorization toward output-oriented, multimodal, and discourse-integrated pedagogies. Yet, productive vocabulary knowledge remains underdeveloped due to gaps in instructional design and contextual alignment. Existing literature supports the use of multimodal tasks and discourse-based scaffolding, but few studies have integrated these under a unified and validated module. This study addresses this gap by proposing a structured, theory-driven vocabulary learning module grounded in SMDM, with instructional delivery based on GRR, and a clear focus on improving learners' productive vocabulary outcomes.

III. THEORETICAL FRAMEWORK

Building on the preceding literature, this study integrates Multimodal Literacy Framework (Jewitt, 2008; Kress & Van Leeuwen, 2001) and Discourse-Based Instruction (Gee, 2008; Hyland, 2014) as the two foundational theoretical pillars guiding the design of the vocabulary learning module. These theories inform content selection, task typology, and instructional strategies by emphasizing meaning-making through multiple semiotic modes and the contextualized use of language in authentic discourse (Coxhead, 2025; Mayer, 2005; Nation, 2006).

To translate these theoretical insights into classroom practice, this study employs the Gradual Release of Responsibility (GRR) model (Fisher & Frey, 2021) as the core instructional delivery framework. Widely operationalized as the "I do, we do, you do" model, GRR scaffolds students' transition from teacher-led modeling to peer collaboration and independent language production. It enables systematic support for learners throughout input, guided practice, and output stages—aligning well with the goals of productive vocabulary instruction.

To ensure that the module design is pedagogically sound, systematically structured, and culturally responsive, the Sidek Module Development Model (SMDM) (Madihie & Noah, 2013; Zainudin et al., 2021) was employed as the overarching instructional design framework. This model provides a stepwise, theory-informed process for module creation, covering goal setting, learner analysis, instructional content design, strategy selection, logistics, and formative validation.

The module is specifically designed to promote productive vocabulary knowledge—the ability to accurately retrieve and use target lexical items in contextually appropriate speech and writing (Laufer, 2005; Nation, 2006). Productive vocabulary is also the primary outcome variable in the module's instructional design and planned evaluation.

Unlike previous interventions that adopt isolated theoretical approaches or overemphasize receptive vocabulary acquisition (Jia, 2024; Teng & Xu, 2022), this study proposes an integrated instructional framework that combines multimodal input-output processing, discourse-rich learning contexts, and scaffolded delivery using the GRR model. This triangulated design offers both theoretical rigor and pedagogical feasibility for enhancing vocabulary instruction among low-proficiency EFL learners in application-oriented universities.

IV. MODULE DESIGN AND DEVELOPMENT

This vocabulary learning module was developed to address the needs of non-English majors in a Chinese application-oriented university, focusing on enhancing their productive vocabulary knowledge. The design was theoretically grounded in Multimodal Literacy Framework (Kress & Van Leeuwen, 2001) and Discourse-Based Instruction (Gee, 2008), and methodologically guided by two complementary frameworks: the Gradual Release of Responsibility (GRR) model (Fisher & Frey, 2021), which structures the scaffolding of classroom instruction from teacher modeling to independent student performance; and the Sidek Module Development Model (SMDM) (Madihie & Noah, 2013).

In the present study, only Phase I (the first nine steps) of the SMDM was implemented and reported. This decision was made to allow for a focused and in-depth presentation of the module's theoretical grounding, content design, and instructional sequencing without the constraints of empirical data reporting. Phase II (validation and refinement) will be addressed in a separate follow-up study. Together, these frameworks ensured that the module was pedagogically coherent, theoretically grounded, and practically feasible. Following are the nine steps of Phase I.

Step 1. Aim of Module Setting

The core objective was to improve students' productive vocabulary knowledge—i.e., their ability to recall and use target vocabulary in contextually appropriate speaking and writing tasks.

Step 2. Theory and Target Group Identification

The design integrated Multimodal Literacy Framework and Discourse-Based Instruction as the core theoretical foundations. The target group consisted of first-year non-English major students at a Chinese application-oriented university.

Step 3. Needs Analysis

Recent studies confirm that Chinese university EFL learners often have strong receptive vocabulary knowledge but struggle with productive use in speaking and writing (Lai & Wang, 2025; Wan et al., 2025). Limited productive vocabulary has been shown to hinder accuracy and fluency in authentic communication (Min & Sukying, 2024). Consistent with these findings, the researcher's decade of teaching experience in applied universities indicates that fewer than 20% of students can use newly learned vocabulary accurately in complete sentences during class, with most relying on L1 translation or single-word responses. While receptive recognition is generally adequate, spontaneous and accurate productive use is rare. These observations stem from regular teaching practice, involved no personal data, and therefore required no formal ethical approval. This convergence of published evidence and classroom experience highlights the need for a theory-driven intervention to strengthen productive vocabulary use.

Step 4. Goal Setting

Specific, measurable goals were developed to improve productive use of vocabulary in communication-oriented tasks.

Step 5. Content Selection

Vocabulary items were drawn from first-semester College English textbooks and verified against COCA and BNC corpora for frequency and authenticity. Items relevant to workplace and daily communication were prioritized.

Step 6. Strategy Selection

Input strategies included the use of images, videos, and audio to introduce vocabulary. Output tasks included infographics, short videos, and storytelling. Instructional strategies were aligned with Cognitive Load Theory and Dual Coding Theory (Paivio, 2014) to promote meaningful retention.

Additionally, each activity was structured using the Gradual Release of Responsibility (GRR) model (Fisher & Frey, 2021), which supports scaffolded learning through three stages: *I do* (teacher modeling), *We do* (guided practice), and *You do* (independent application). This model ensured that students received adequate support when transitioning from multimodal vocabulary input to independent language production tasks.

Step 7. Logistics Planning

The module was integrated into a 16-week semester, with weekly 4-hour sessions combining vocabulary input, multimodal tasks, and formative assessment. Digital platforms such as Chaoxing Learning Platform were used for homework and projects.

Step 8. Media and Technology Selection

Tools such as PowerPoint and Quizlet were selected to support both input delivery and student production, promoting multimodal literacy and creative expression.

Step 9. Module Combination Process

All instructional elements were compiled into a sequenced module, organized by thematic weekly units aligned with the curriculum. Activities within each unit followed a structured delivery pattern based on the GRR instructional sequence, progressing from teacher-led demonstrations to collaborative work and then to independent application. This allowed students to internalize vocabulary knowledge progressively while engaging in multimodal and discourse-based tasks.

V. DESIGN OUTPUT AND THEORETICAL JUSTIFICATION

A. Module Architecture Overview

This multimodal vocabulary learning module is systematically designed to enhance the productive vocabulary knowledge of first-year non-English majors at Chinese application-oriented universities. The module is grounded in Multimodal Literacy Framework (Jewitt, 2008; Kress & Van Leeuwen, 2001), Discourse-Based Instruction (Gee, 2008; Hyland, 2014), and the Gradual Release of Responsibility (GRR) Model (Fisher & Frey, 2021), ensuring pedagogical alignment with learners' cognitive, linguistic, and affective needs.

The structure of the module follows two main dimensions:

1) Horizontal Structure (Lesson Sequence):

The module comprises 5 instructional units, each delivered over 2 class sessions, with 4 interconnected activities per session, totaling 8 activities per unit. These activities are organized in a spiraling and progressive sequence that supports learners' transition from perceptual input to creative output. The design ensures cognitive scaffolding and gradual transfer of responsibility from teacher to learner.

2) Vertical Structure (GRR Pedagogical Scaffolding):

Each activity is framed within the three stages of the GRR model:

I do: The teacher models the target vocabulary through pronunciation, spelling, or usage.

We do: The teacher guides the class through structured practice, emphasizing joint meaning construction and formative feedback.

You do: Students independently complete output tasks such as sentence creation, oral production, or multimodal projects.

Table 1 summarizes the core design elements of the module:

TABLE 1
MODULE DESIGN OVERVIEW

Dimension	Description
Instructional Length	4 units × 2 lessons = 8 lessons
Tasks per Session	4 tasks per session, 8 tasks per unit
Target Learners	First-year non-English majors in Chinese application-oriented universities
Learning Focus	Productive Vocabulary Knowledge (PVK)
Theoretical Foundations	Multimodal Literacy Framework, Discourse-Based Instruction, GRR
Delivery Format	In-class instruction + digital homework
Technological Tools	PowerPoint, Quizlet, Chaoxing Learning Platform, etc.

This architecture prioritizes rich language context, modal integration, and progressive cognitive engagement, with the GRR model serving as a backbone for both instructional delivery and teacher behavior observation. It aims to foster meaningful, transferable vocabulary production in EFL contexts specific to the needs of applied university settings in China.

B. Sample Tasks and Theoretical Mapping

The multimodal vocabulary learning module comprises eight interrelated activities per unit, each purposefully designed to activate various modes of learning and align with specific theoretical underpinnings. Each unit is delivered over two lessons, with four activities implemented per lesson to ensure focused progression and cognitive manageability. Table 2 below presents representative tasks from Unit 1 and maps each to its corresponding theoretical foundation and instructional function within the Gradual Release of Responsibility (GRR) framework.

Each task is designed not in isolation, but as part of a cohesive instructional sequence. For instance, vocabulary introduced through multimodal input in *Activity 1* is reinforced through collaborative practice in *Activity 2 and 3*, and then further internalized through independent creation in *Homework 1*. This flow ensures consistent application of the “I do – We do – You do” structure while scaffolding learners toward autonomous vocabulary use.

TABLE 2
SAMPLE TASKS WITH GRR PHASES AND THEORETICAL FOUNDATIONS OF UNIT

	Activity Name	Modality Used	Objective	Theoretical Support	Materials	GRR Implementation
Lesson 1	1 Vocabulary Presentation and Pronunciation Practice	Visual & Auditory	Memorize meaning and pronunciation of vocabulary through multi-sensory input	Multimodal Literacy Framework	Slides (vocabulary, images, pronunciation)	I do: Teacher models pronunciation with visuals; We do: Students repeat with guidance; You do: Students pronounce and write target words independently
	2 Listening and Sentence Writing	Auditory & Textual	Understand and apply vocabulary via listening and writing	Discourse-Based Instruction	Listening quiz, recordings	I do: Teacher demonstrates how to listen for and note key vocabulary. We do: Teacher and students co-construct a model sentence. You do: Students write their own sentences based on the listening task.
	3 Vocabulary Puzzle Game	Kinesthetic, Visual & Textual	Reinforce vocabulary through kinesthetic group activity	Multimodal Literacy Framework	Vocabulary cards (physical or digital), Quizlet match activity, image/context cards	I do: Teacher demonstrates matching game using Quizlet; We do: Students complete initial rounds in pairs using Quizlet; You do: Each student completes the match individually and reflects on incorrect responses.
	4 Homework 1: Vocabulary Log	Visual & Textual	Reinforce understanding and practical application of the vocabulary.	Multimodal Literacy Framework; Discourse-Based Instruction	Storyboard templates or digital creation tools; Super Star platform	I do: Teacher shows examples of completed logs and explains how to use images + sentences. We do: Teacher helps students brainstorm example entries. You do: Students complete vocabulary logs independently with sentence + image.
Lesson 2	5 Vocabulary Relay	Visual, Textual & Kinesthetic	Enhance vocabulary recall and engagement via physical spelling game	Multimodal Literacy Framework	Slides (images and definitions)	I do: Teacher demonstrates the gameplay. We do: Teams practice the relay in low-stakes rounds. You do: Students participate independently and retrieve words under timed conditions.
	6 Discourse-Based Instruction	Auditory & Textual	Understand vocabulary in discourse-based real-world contexts	Discourse-Based Instruction	Articles with vocabulary, audio & text	I do: Teacher guides reading and explains vocabulary in context. We do: Teacher and students annotate and discuss key words in the text. You do: Students summarize or write reflections using the new words.
	7 Group Discussion and Story Creation	Visual, Auditory & Textual	Integrate vocabulary in collaborative and creative story tasks	Multimodal Literacy Framework	Story worksheet, rubric, digital tools	I do: Teacher models how to turn vocabulary into narrative. We do: Students collaboratively write a short story using key vocabulary. You do: Each student contributes a scene or sentence independently.
	8 Homework 2: Storyboard or Digital Story Creation	Visual, Auditory & Textual	Consolidate understanding and application of vocabulary in a real-world context and demonstrate creative achievements.	Multimodal Literacy Framework; Discourse-Based Instruction	Storyboard templates or digital creation tools; Super Star platform	I do: Teacher explains expectations and shows story examples. We do: Teacher helps students plan out storyboards. You do: Students independently complete digital or drawn story using vocabulary.

Moreover, the tasks emphasize productive vocabulary knowledge (PVK)—the ability to generate accurate and contextually appropriate vocabulary in speaking and writing. This is particularly essential for low-proficiency EFL learners in applied universities, who often lack sufficient opportunities for output-based practice.

By grounding each activity in both theoretical rationale and instructional progression, the module ensures pedagogical coherence, learner engagement, and replicability in similar educational contexts.

C. Pedagogical Innovation and Transferability

This vocabulary learning module demonstrates significant pedagogical innovation within the context of English instruction for non-English majors at Chinese application-oriented universities. Departing from traditional vocabulary teaching practices—such as rote memorization, word lists, and isolated translation exercises—this module is grounded in Multimodal Literacy Framework and Discourse-Based Instruction, and is implemented through the Gradual Release of Responsibility (GRR) model, with the primary aim of enhancing students' productive vocabulary knowledge, particularly their ability to use target words accurately and appropriately in real-world contexts.

A key innovation of the module lies in its intentional multimodal integration. Each activity is designed to engage learners through two to three sensory modalities (e.g., visual + auditory, kinesthetic + textual, etc.), thereby activating multiple cognitive channels to deepen comprehension, support memory retention, and promote meaningful language use. Moreover, the cognitive sequence of each activity progresses from input to output, and from low- to high-order thinking. This is achieved through the GRR framework, which guides students through three stages: teacher modeling (I do), guided practice (We do), and independent application (You do). Such a scaffolding mechanism ensures balanced instructional support while fostering learner autonomy and communicative competence.

Additionally, the use of the Sidek Module Development Model (SMDM) ensures that the entire design process—from objective formulation and task design to instructional strategy and evaluation—is systematic, theoretically coherent, and culturally responsive.

In terms of transferability, although this module is tailored for Chinese application-oriented universities, its design principles—multimodal scaffolding, discourse embedding, structured progression, and task alignment—are highly generalizable. With flexible tools such as Quizlet, educators across different contexts can adapt the module to suit their curricular and cultural needs.

In sum, this module offers an innovative, theory-driven alternative to decontextualized, teacher-centered vocabulary instruction. It provides a replicable, adaptable, and scalable instructional framework for EFL educators aiming to shift toward output-focused, learner-engaged, and multimodally supported vocabulary teaching.

D. Comparison With Existing Practices

As reviewed earlier, vocabulary instruction in Chinese universities has been shifting toward greater contextualization and communicative relevance, with increasing adoption of multimodal strategies and discourse-based activities to enhance learner engagement (Pan, 2023). These innovations reflect a growing recognition of the need to connect vocabulary learning with authentic language use in academic and professional contexts (Jia, 2024). However, classroom implementations often remain fragmented, short-term, and lacking in sustained scaffolding or systematic theoretical integration (Argyriou & Tapsis, 2025; Huang, 2025; Muhammad Nur Iman & Nurasia Natsir, 2025). While such practices may introduce diverse techniques, they frequently operate without a coherent instructional design framework, making it difficult to ensure continuity, contextual relevance, and replicability across lessons and courses.

In contrast, the module developed in this study is explicitly designed to address these limitations. First, it shifts the instructional focus from receptive recognition to productive vocabulary development through output-oriented, multimodal, and discourse-based tasks—such as storytelling, infographic creation, video production, and collaborative discussions—that require students to use target vocabulary meaningfully in authentic contexts. Second, its instructional delivery follows the Gradual Release of Responsibility (GRR) model (Fisher & Frey, 2021), ensuring a scaffolded learning process from teacher modeling to guided practice and independent production. Finally, the module's development adheres to the Sidek Module Development Model (SMDM) (Madihie & Noah, 2013; Zainudin et al., 2021), which provides a structured and validated framework to ensure theoretical alignment, cultural contextualization, and instructional feasibility. This integration of theoretically grounded pedagogy with systematic design stands in marked contrast to the fragmented and input-heavy practices prevalent in many existing EFL classrooms in Chinese application-oriented universities.

E. Discussion of Theoretical Coherence

The strength of this vocabulary learning module lies in its deliberate integration of multiple theoretical frameworks to ensure both pedagogical robustness and contextual adaptability. Each design decision—from content sequencing to modality alignment and classroom task progression—was guided by a triangulated theoretical foundation that includes Multimodal Literacy Framework, Discourse-Based Instruction, and the Gradual Release of Responsibility (GRR) model.

Multimodal Literacy Framework (Jewitt, 2008; Kress & Van Leeuwen, 2001) provided the epistemological grounding for integrating visual, auditory, kinesthetic, and textual modalities in each activity. Rather than treating language in isolation, this perspective allowed vocabulary instruction to be situated within meaning-making practices across multiple semiotic channels. For instance, activities such as storyboard creation and vocabulary puzzles required learners to engage with visual and kinesthetic input while simultaneously producing linguistic output, thus deepening cognitive processing and memory encoding (Mayer, 2005).

Discourse-Based Instruction (Gee, 2008; Hyland, 2014) informed the contextualization of vocabulary tasks within authentic communicative situations. Unlike traditional approaches that isolate vocabulary lists from meaningful use, this module embedded lexical items within narrative texts, dialogues, and student-created discourse. This ensured that vocabulary learning occurred not only at the lexical level, but also in the broader context of pragmatics and genre-based competence (Webb & Nation, 2017).

To operationalize these theories at the classroom delivery level, the GRR model (Fisher & Frey, 2021) was adopted as a sequencing mechanism. The “I do–We do–You do” scaffold guided the progression of each activity, beginning with teacher modeling, followed by guided practice, and culminating in independent student performance. This model provided a clear, research-backed structure for balancing support and autonomy, which is particularly crucial for low-proficiency learners in application-oriented universities (Fisher & Frey, 2021). It also ensured that each multimodal task was not implemented as a discrete activity, but as part of a sequenced instructional arc that aligns with the cognitive development of the learner.

Moreover, the entire module design process was structured using the Sidek Module Development Model (SMDM) (Madihie & Noah, 2013), which served as a meta-framework ensuring internal consistency and theoretical alignment across all stages—from needs analysis and goal setting to content design and formative assessment. This layered integration of theories and models resulted in a module that is not only pedagogically coherent but also scalable, transferable, and responsive to the evolving goals of applied English instruction in higher education.

In sum, the module reflects a synergistic application of theoretical perspectives, translating complex educational principles into practical, observable, and replicable classroom strategies. This coherence strengthens the module's potential for impact and serves as a model for theory-driven instructional innovation in the field of English as a Foreign Language (EFL).

F. Ethical Approval and Informed Consent

This study did not involve the collection of personal or sensitive data from human participants. Therefore, ethical approval and informed consent were not required. However, the module design and description were developed in accordance with professional ethical standards in educational research.

VI. CONCLUSION

The development of this theory-driven multimodal vocabulary learning module carries meaningful pedagogical and theoretical implications for EFL instruction in application-oriented university contexts. By integrating the principles of Multimodal Literacy Framework and Discourse-Based Instruction with the Gradual Release of Responsibility (GRR) framework, the module offers a structured yet flexible blueprint for enhancing learners' productive vocabulary knowledge in ways that are both cognitively engaging and contextually relevant. Unlike traditional vocabulary instruction that often prioritizes recognition and rote memorization, this module underscores the importance of language output, multimodal representation, and scaffolded instructional delivery. Such a shift aligns with global trends in competency-based language education and responds to the increasing demands for communication-oriented English teaching in Chinese higher education.

The design also has the potential to contribute to the field of instructional design. Drawing on the Sidek Module Development Model (SMDM), the module was developed through a sequenced, evidence-informed process intended to bridge theory and classroom implementation. The planned use of GRR as a delivery framework is expected to ensure that each learning activity incorporates clear modeling, guided practice, and opportunities for independent use, thereby supporting learners' gradual transition from supported input to autonomous output. This structure may be particularly beneficial for low-proficiency learners, who often lack the confidence or competence to apply new vocabulary independently without systematic scaffolding. While the present study does not report on classroom implementation, the model holds strong potential for adaptation across a range of EFL teaching environments, especially those aiming to foster deeper vocabulary engagement and long-term retention.

While the current study focuses on module construction rather than empirical testing, the design itself serves as a foundational contribution that may be further validated and refined through future classroom-based implementation and research. The modular structure and task typology can be readily adopted or adapted by practitioners in other contexts. Moreover, the theoretical integration proposed herein invites further scholarly exploration of how multimodal input, discourse-rich contexts, and scaffolded instruction may interact to support different learner profiles and vocabulary outcomes. Future research may examine the long-term effects of such modules on learner autonomy, motivation, and communicative competence, as well as explore technological extensions such as digital platforms or AI-assisted multimodal tasks to scale the implementation. In sum, the present module not only addresses an immediate pedagogical gap, but also lays the groundwork for a replicable, theory-informed model of vocabulary instruction that can evolve alongside emerging educational needs and technological advancements.

To conclude, this study presents the design and theoretical foundation of a structured multimodal vocabulary learning module aimed at enhancing productive vocabulary knowledge among non-English majors at a Chinese application-oriented university. Grounded in Multimodal Literacy Framework and Discourse-Based Instruction, and scaffolded by the Gradual Release of Responsibility (GRR) model, the module was systematically constructed through the Sidek Module Development Model (SMDM). This integrated approach allowed for the alignment of content, pedagogical strategies, and task design to support meaningful vocabulary use in both spoken and written contexts.

Unlike fragmented instructional practices or theory-light applications, this module demonstrates how a coherent and replicable framework can be developed to promote vocabulary learning through multimodal input and structured output. The incorporation of GRR ensures that each activity transitions from teacher modeling to collaborative practice and

finally to independent application, thus addressing the needs of low-proficiency learners for support, practice, and autonomy. The resulting module not only reflects sound pedagogical principles but also responds to the practical realities and policy expectations of China's application-oriented higher education system.

Although this article does not include empirical findings, the completed design provides a strong foundation for future validation studies. Follow-up research may examine the implementation of the module in real classroom contexts, evaluate learner engagement and outcomes, and further refine its components based on empirical data. Ultimately, this module offers not only an innovative contribution to vocabulary instruction in EFL contexts but also a scalable and theoretically grounded model for designing integrated language learning materials.

REFERENCES

- [1] Aedo, P., & Millafilo, C. (2022). Increasing vocabulary acquisition and retention in EFL young learners through the use of multimodal texts (memes). *Colombian Applied Linguistics Journal*, 24(2), 251–269. <https://doi.org/10.14483/22487085.18312>
- [2] Al-Jarf, R. (2023). Multimodal teaching and learning in the EFL college classroom. *Journal of English Language Teaching and Applied Linguistics*, 6(4), 68-75. <http://dx.doi.org/10.32996/jeltal.2024.6.4.7>
- [3] Argyriou, M., & Tapsis, N. (2025). Reframing multimodality in contemporary education: A critical review of research patterns and pedagogical shifts. *European Journal of Open Education and E-Learning Studies*, 10(3). <https://doi.org/10.46827/ejoe.v10i3.6106>
- [4] Bai, Yijia. (2023). English vocabulary learning strategies of undergraduates from a multimodal perspective. *Advances in Education*, 13(06), 3229–3233. <https://doi.org/10.12677/AE.2023.136509>
- [5] Bansong, K., Poopatwiboon, S., & Sukying, A. (2023). The effects of multimodal teaching on English vocabulary knowledge of Thai primary school students. *Journal of Education and Learning*, 12(6), 46. <https://doi.org/10.5539/jel.v12n6p46>
- [6] Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. Springer US. <https://doi.org/10.1007/978-0-387-09506-6>
- [7] Caro, K., & Mendinueta, N. R. (2017). Lexis, lexical competence and lexical knowledge: a review. *Journal of Language Teaching & Research*, 8(2).
- [8] Coxhead, A. (2025). Vocabulary and English for Specific Purposes. In *The Handbook of English for Specific Purposes* (pp. 107–123). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119985068.ch6>
- [9] Fisher, D., & Frey, N. (2021). *Better Learning Through Structured Teaching: A Framework for the Gradual Release of Responsibility*. ASCD.
- [10] Fu, C. (2024). Discussion on the reform and practice of college English teaching in application-oriented undergraduate institutions under the background of 'golden course'. *Journal of Educational Theory and Practice*, 1(4), 5–9. <https://doi.org/10.62177/jetp.v1i4.89>
- [11] Gee, J. P. (2008). *Social linguistics and literacies: Ideology in discourses* (3rd ed.). Routledge.
- [12] Huang, W. (2025). Unveiling the power of multi-modality in literacy learning: Navigating literacy education in Chinese cultural contexts. *International Journal of Education and Research*, 13(6).
- [13] Hyland, K. (2014). English for Academic Purposes. In *The Routledge Companion to English Studies*. Routledge.
- [14] Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, 32(1), 241–267. <https://doi.org/10.3102/0091732X07310586>
- [15] Jia, W. (2024). Digital game-based learning's effectiveness on EFL learners' receptive and productive vocabulary knowledge. *Language Learning & Technology*, 28(1).
- [16] Jiang, Y. (2020). Study on the development of critical thinking in college English teaching in application-oriented university. In *7th International Conference on Education, Language, Art and Inter-cultural Communication (ICELAIC 2020)* (pp. 13-18). Atlantis Press. <https://doi.org/10.2991/assehr.k.201215.293>
- [17] Kress, G. R., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Hodder Education.
- [18] Lai, C., & Wang, Q. (2025). Profiles of media use purpose in informal digital learning of English and their association with vocabulary knowledge. *Computer Assisted Language Learning*, 1–31. <https://doi.org/10.1080/09588221.2025.2501705>
- [19] Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, 5(1), 223–250. <https://doi.org/10.1075/eurosla.5.11lau>
- [20] Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 36–55. <https://doi.org/10.1191/026553299672614616>
- [21] Liang, Y. (2021). Research on current situation and countermeasures of English teaching in independent colleges—Based on application-oriented universities transformation. In *Proceedings of the 2021 4th International Conference on Humanities Education and Social Sciences (ICHESS 2021)*. Atlantis Press. <https://doi.org/10.2991/assehr.k.211220.005>
- [22] Madihie, A., & Noah, S. M. (2013). An Application of the Sidek Module Development in Rebt Counseling Intervention Module Design for Orphans. *Procedia - Social and Behavioral Sciences*, 84, 1481–1491. <https://doi.org/10.1016/j.sbspro.2013.06.777>
- [23] Mayer, R. E. (Ed.). (2005). *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press.
- [24] Meng, Q., Liu, W., & Lv, Y. (2023). Research on the teaching reform of translation online course for business English majors in application-oriented universities—A case study of S university. *SHS Web of Conferences*, 159, 01005. <https://doi.org/10.1051/shsconf/202315901005>
- [25] Muhammad Nur Iman, & Nurasia Natsir. (2025). Translanguaging pedagogies in digital learning environments: Fostering multilingual competence among generation Z learners. *Proceeding of The International Conference of Inovation, Science, Technology, Education, Children, and Health*, 5(1), 57–74. <https://doi.org/10.62951/icitech.v5i1.268>
- [26] Nation, I. S. P. (2006). *Learning Vocabulary in Another Language* (8th ed.). Cambridge University Press.
- [27] Paivio, A. (2014). *Mind and Its Evolution: A Dual Coding Theoretical Approach*. Psychology Press. <https://doi.org/10.4324/9781315785233>

- [28] Pan, R. (2023). The role of receptive/orthographic vocabulary, productive/orthographic vocabulary, productive/phonological vocabulary and depth of vocabulary in predicting reading-to-write performance. *Journal of Psycholinguistic Research*, 52. <https://doi.org/10.1007/s10936-023-10027-8>
- [29] Qiann, L., Mai, J., & Marie, P. A. (2025). Discourse-Based English Instructional Model: A Lesson from Pedagogical Discourse Analysis and Communicative Approaches. *Ecolingua: Journal of Linguistics and Language Instruction*, 1(1), 1-13.
- [30] Scandurra, R., Kelly, D., Fusaro, S., Cefalo, R., & Hermannsson, K. (2024). Do employability programmes in higher education improve skills and labour market outcomes? A systematic review of academic literature. *Studies in Higher Education*, 49(8), 1381–1396. <https://doi.org/10.1080/03075079.2023.2265425>
- [31] Schmitt, N., Nation, P., & Kremmel, B. (2020). Moving the field of vocabulary assessment forward: The need for more rigorous test development and validation. *Language Teaching*, 53(1), 109–120. <https://doi.org/10.1017/S0261444819000326>
- [32] Simmons, S. E. (2012). *The Origins and Descriptions of ADDIE*. <http://www.shelbyesimmons.me/uploads/1/0/4/1/10414302/simmonsseaddiepaper.pdf>
- [33] Teng, M. F. (2023). The effectiveness of multimedia input on vocabulary learning and retention. *Innovation in Language Learning and Teaching*, 17(3), 738–754. <https://doi.org/10.1080/17501229.2022.2131791>
- [34] Teng, M., & Xu, J. (2022). Pushing vocabulary knowledge from receptive to productive mastery: Effects of task type and repetition frequency. *Language Teaching Research*, 29(2), 588-606. <https://doi.org/10.1177/13621688221077028>
- [35] Wakasa, D. B. (2024). Discourse Analysis and its Role in Enhancing ESL Learners Reading Comprehension Competencies: A Literature Review. *Journal of Frontiers in Humanities and Social Sciences*, 2(2), 63-75. <https://doi.org/10.69897/jofhscs.v2i2.118>
- [36] Wan, C., Abdullah, A. N., Bolong, J., Habil, H., & Nimehchisalem, V. (2025). Effect of baicizhan application on English vocabulary knowledge on non-English major university students. *Open Journal of Modern Linguistics*, 15(02), 162–177. <https://doi.org/10.4236/ojml.2025.152012>
- [37] Webb, S., & Nation, P. (2017). *How Vocabulary is Learned*. Oxford University Press.
- [38] Xu, M., & Zhang, C. (2021). A study on curriculum reform of business English in application-oriented universities in the new era. In *2021 2nd International Conference on Modern Education Management, Innovation and Entrepreneurship and Social Science (MEMIESS 2021)* (pp. 251-255). Atlantis Press.
- [39] Zainudin, A. Z., Yunus, N. M., Zakaria, S. R. A., & Mohsin, A. (2021). Design for sustainability integration in education. In *Design for Sustainability* (pp. 139–169). Elsevier. <https://doi.org/10.1016/B978-0-12-819482-9.00002-2>
- [40] Zhao, X., Wang, L., & Chen, L. (2024). Industry and Education Integration for Chinese University of Applied Science and Technology, Based on the Triple Helix Theory. In *3rd International Conference on Education, Language and Art (ICELA 2023)* (pp. 316-328). Atlantis Press. https://doi.org/10.2991/978-2-38476-214-9_38

Meng Yaqiong was born in China. She holds an M.A. in TESOL from the Education University of Hong Kong and is currently pursuing a Ph.D. in TESL at the Faculty of Education, Universiti Kebangsaan Malaysia (UKM). Her research interests include multimodal instructional design, productive vocabulary learning and teaching, and the role of foreign language enjoyment in technology-enhanced classrooms. ORCID iD: <https://orcid.org/0009-0004-2920-0054> Email: p125436@siswa.ukm.edu.my

Supyan Hussin is Professor of Applied Linguistics and Senior Research Fellow at the Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia (UKM). He earned his Ph.D. in educational policy studies with a focus on language education from the University of Illinois at Urbana-Champaign, USA, an M.A. in applied linguistics from the University of Essex, UK, and a B.Ed. in TESL from UKM. His research interests include technology-enhanced language learning, multimodal pedagogy, higher-education language policy, and digital literacy. He has led multiple nationally funded projects on mobile-assisted vocabulary learning and has published widely on integrating multimodal resources into ESL classrooms. ORCID iD: <https://orcid.org/0000-0003-0544-0649> Email: supyan@ukm.edu.my

Harwati Hashim is Professor of Educational Technology at the Faculty of Education, Universiti Kebangsaan Malaysia (UKM). She received her Ph.D. in educational technology from UKM, an M.Ed. in instructional technology from UKM, and a B.Ed. in TESL from UKM. Her research centres on mobile-assisted language learning, multimodal instructional design, learner analytics, and augmented reality for vocabulary acquisition. She has secured several national grants exploring the design and impact of multimodal learning modules among low-proficiency tertiary learners and has published extensively in journals and edited volumes on technology-mediated language education. ORCID iD: <https://orcid.org/0000-0002-8817-427X> Email: harwati@ukm.edu.my