WHIOCE PUBLISHING PTE. LTD.
PROVIDING
FIRST-CLASS SCIENTIFIC INFORMATION
FOR TOP SCHOLARS

2025 Volume 3, Issue 11
Contemporary Education Frontiers
ISSN:3029-1879 (Online)
ISSN: 3029-1860 (Print)

Research on Intelligent Teaching System of International Chinese Vocabulary Integrating Knowledge Graphs and Large Language Models

Jiayi Lin*
Hainan Vocational University of Science and Technology, Haikou 571126, Hainan, China

Author to whom correspondence should be addressed.

Abstract: With the rapid advancement of artificial intelligence technology, its applications in the education sector have become increasingly profound. As a vital bridge for cross-cultural communication, international Chinese language teaching faces urgent demands for innovative teaching methodologies and efficiency improvements, particularly in core vocabulary instruction. Traditional vocabulary teaching methods often struggle to address learners' individualized needs, the complex semantic relationships between words, and their rich cultural connotations. This paper explores a novel intelligent teaching system for international Chinese vocabulary that integrates knowledge graphs with large language models. Leveraging the powerful natural language understanding, generation, and interaction capabilities of large language models, the system provides more dynamic and user-friendly interfaces for knowledge graphs. The study systematically outlines the technical foundations and core mechanisms of knowledge graph-large language model integration, designs a system architecture tailored for international Chinese vocabulary teaching, and details the implementation approaches for key functional modules, including personalized learning path planning, contextualized vocabulary analysis, intelligent Q&A, and assessment generation. Finally, the paper examines multi-dimensional evaluation methods for the system and analyzes current technical challenges along with future development directions.

Keywords: Knowledge graph; Large language model; International Chinese education; Vocabulary teaching; Intelligent teaching system

Online publication: Dcecmber 20, 2025

1. Introduction

1.1. Research background and significance

International Chinese education serves as a crucial pathway for promoting Chinese culture globally and facilitating cultural exchange and mutual learning. As the cornerstone of language, the breadth and depth of vocabulary mastery directly determine learners' communicative competence. However, international Chinese vocabulary acquisition faces multiple challenges as follows:

(1) The phenomena of "multiple meanings for a single character" and "multiple meanings for a single word" are widespread, with complex relationships such as synonymy, antonymy, and hierarchical relationships among words;
(2) A significant portion of vocabulary carries rich cultural contexts and pragmatic situations, making rote memorization without contextual awareness largely ineffective;
(3) Learners come from diverse backgrounds, exhibiting substantial variations in learning pace, difficulties, and cognitive patterns, making the traditional "one-size-fits-all" teaching approach inadequate for personalized instruction.

In recent years, artificial intelligence technologies represented by large language models (LLMs) have demonstrated revolutionary potential in language processing and content generation. However, while LLMs enable fluent and natural interactions, they also exhibit inherent limitations such as knowledge "illusion," factual errors, and insufficient reasoning capabilities [1]. On the other hand, knowledge graphs (KGs), as structured semantic networks, accurately represent concepts and their interrelationships, offering advantages of high knowledge accuracy, strong logical coherence, and excellent interpretability. Therefore, combining the structured knowledge strengths of knowledge graphs with the linguistic capabilities of large language models to develop a teaching system that is both "vastly knowledgeable" and "precise," as well as "intelligent" and "reliable," holds significant theoretical and practical importance for addressing the aforementioned challenges in international Chinese vocabulary instruction.

1.2. Current research status

Currently, research on the integration of knowledge graphs and large language models has become a hot topic in both academia and industry. Extensive studies focus on the methodologies, technical approaches, and general applications of this integration. For instance, in the education sector, scholars have explored combining the two to develop advanced educational QA systems by leveraging knowledge graphs to enhance query-answering accuracy and contextual awareness. Additionally, research utilizing multi-agent workflows to learn knowledge component graph structures has demonstrated their potential for personalized teaching interventions.

However, a systematic review of academic literature and project reports from 2022 to 2025 reveals that despite extensive technical discussions, concrete system implementations specifically tailored for the field of "intelligent teaching of international Chinese vocabulary" remain exceedingly scarce. Most existing research remains at the level of theoretical exploration, general framework design, or preliminary applications in other languages (e.g., Japanese translation instruction) or disciplines (e.g., environmental science). Few studies have delved into designing teaching platforms for international Chinese learners that integrate vocabulary knowledge graphs with large language models (LLMs), featuring comprehensive system architectures, data processing workflows, and detailed algorithmic specifications. Additionally, standardized evaluation datasets and robust assessment methodologies for such systems have yet to be established, representing a significant gap in current research [2].

1.3. Research content of this paper

In light of this, this study aims to address the aforementioned gaps by systematically investigating the application of integrating knowledge graphs with large language models in intelligent international Chinese vocabulary teaching systems. The main research components include as outlined:
(1) Technical foundation analysis. providing an in-depth examination of the respective advantages and limitations of knowledge graphs and large language models, along with elucidating the complementary principles and mainstream technical approaches for their integration;
(2) System architecture design, proposing a hierarchical, decoupled overall architecture for international Chinese vocabulary instruction with clearly defined functions for each core component;
(3) Core functionality implementation, detailed design of key system modules such as personalized learning path planning, contextual vocabulary analysis, intelligent Q&A, and assessment generation;
(4) Evaluation methodology, developing a dual assessment framework encompassing technical performance evaluation

and teaching effectiveness assessment, accompanied by specific user experiment designs;
(5) Challenges and prospects, summarizing major current challenges in technological application and outlining future development trends.

2. Technical foundations for the integration of knowledge graphs and large language models

2.1. Advantages and limitations of large language models

Large language models (such as the GPT series and Gemini) have developed powerful natural language processing capabilities through pre-training on vast amounts of text data. Their advantages are primarily manifested in three aspects:
(1) Exceptional language generation and comprehension abilities, enabling them to produce fluent, coherent, and grammatically correct text while deeply understanding contextual nuances;
(2) Robust generalization and zero/low-sample learning capabilities, allowing them to handle tasks and problems not encountered in the training data;
(3) Natural interaction experiences that support multi-turn dialogue and facilitate conversational engagement with users in an anthropomorphic manner [3].

However, its limitations have become increasingly apparent. For instance:
(1) The knowledge illusion (Hallucination), which may generate information that appears reasonable but contradicts reality;
(2) Delayed knowledge updates, as the refreshment of its internal knowledge base relies on model retraining, a process that is costly and prone to issues of timeliness;
(3) Limited reasoning capabilities, with inconsistent performance on problems requiring complex logical reasoning and multi-step derivations;
(4) Poor explainability, as the decision-making process is difficult to trace and the origin of answers cannot be clearly explained.

2.2. Characteristics of knowledge graphs

A knowledge graph is a technology that uses graph structures to model relationships between knowledge and entities. Its core advantages include as follows:
(1) Structured and precise representation, storing knowledge in "entity-relation-entity" triples that ensure clear information and minimize ambiguity;
(2) Robust logical reasoning capabilities, enabling symbolic logical inference based on the graph structure to uncover implicit knowledge;
(3) High interpretability and reliability, with traceable knowledge sources and transparent reasoning paths.

Its limitations include as follows:
(1) The high construction cost, as building a high-quality knowledge graph requires substantial manual effort and specialized expertise;
(2) Limited knowledge coverage, typically confined to specific domains and unable to encompass all knowledge in an open world;
(3) Weak natural language interaction capabilities, lacking the ability to engage in fluent dialogue with users.

2.3. Complementarity and core pattern of integration

LLM and KG complement each other perfectly in terms of capabilities. LLM serves as a "language bridge" connecting users with structured knowledge, while KG provides LLM with "fact anchors." Their integration aims to achieve a synergistic effect where "1+1>2." According to current research, the integration models primarily fall into three categories:

The first category is knowledge graph-enhanced large language models (KG-Enhanced LLMs), which represent the most prevalent integration paradigm today. The core concept involves leveraging knowledge graphs to improve LLM performance. The most typical implementation is Retrieval-Augmented Generation (RAG). In this approach, when an LLM receives a question, the system first retrieves relevant and accurate knowledge fragments (such as entities, relationships, and attributes) from the knowledge graph, then incorporates these fragments as context prompts alongside the original question to guide the LLM in generating more precise and reliable answers. This method effectively reduces hallucinations and integrates the latest external knowledge.

The second category comprises LLM-enhanced knowledge graphs (LLM-Enhanced KGs), which leverage the powerful text comprehension capabilities of large language models to empower the entire lifecycle of knowledge graphs, including automated construction (extracting triples from unstructured text), completion (predicting missing links), and natural language queries (converting users natural language queries into graph query languages such as SPARQL).

The third category comprises Synergistic and Deep Fusion Frameworks, which aim for deeper integration. For instance, during model training, the structural information of knowledge graphs is incorporated into the LLMs parameters, or a collaborative inference framework is designed to enable dynamic interaction and iterative refinement between the LLM and KG during inference, enabling them to jointly address complex problems.

For international Chinese vocabulary teaching systems, the "knowledge graph-enhanced large language model" approach is particularly suitable, as it directly meets the stringent requirements for knowledge accuracy in teaching scenarios [4].

3. System architecture design for international Chinese vocabulary teaching

3.1. Overall architecture

To achieve the aforementioned objectives, this study proposes a modular and scalable system architecture that draws inspiration from the design philosophy of general-purpose educational platforms and employs a microservices architecture to enhance flexibility and maintainability. The system adopts a hierarchical design, consisting of four layers from bottom to top:

The Infrastructure Layer provides fundamental computing, storage, and network resources, deployable on either public or private clouds, leveraging containerization technologies (such as Docker and Kubernetes) for elastic scaling and efficient resource management.

The Data Resource Layer serves as the core data asset of the system, comprising three key components:

(1) The International Chinese Vocabulary Knowledge Graph (KG), which forms the foundation of the systems knowledge base. This graph integrates various node types, including vocabulary, Chinese characters, example sentences, images, audio recordings, and cultural elements, with extensive relationship edges covering synonyms, antonyms, parts of speech, etymology, collocations, and pragmatic contexts, linking characters, words, sentences, and grammatical structures to form an extensive semantic network;

(2) The Large Language Model, which employs a foundational model with outstanding Chinese processing capabilities (evaluated using benchmarks like C-Eval and CMMLU) and is fine-tuned with international Chinese teaching corpora to better suit educational scenarios;

(3) The Learner Model Database stores users basic information, learning histories, vocabulary mastery levels, error records, and interaction logs, providing data support for personalized recommendations.

The Application Service Layer serves as the functional core of the system, comprising a suite of microservices: the Knowledge Graph Management Service, responsible for storing, querying, updating, and reasoning with the Knowledge Graph (KG); the LLM Reasoning Service, providing access interfaces for fine-tuned large language models; the Fusion Orchestration Engine, acting as the systems "brain," which parses user requests, determines the appropriate KG-LLM integration strategy (e.g., RAG), and integrates results from various services; the Teaching Application Service, delivering specific educational features such as learning path planning, intelligent Q&A, and exercise generation; and the User Management and Analysis Service, which manages user information and analyzes learning behaviors.

The User Interaction Layer serves as the front-end interface for user interaction with the system, encompassing Web applications, mobile apps, or chatbots integrated into instant messaging tools.

3.2. Detailed description of core components

3.2.1. International Chinese vocabulary knowledge graph

Constructing this ontology is the key and most challenging aspect of the project. Data sources may include authoritative dictionaries (such as the Modern Chinese Dictionary), standard syllabus vocabulary lists like HSK, open encyclopedias, and specialized corpora. Large Language Models (LLMs)can be employed to perform semi-automatic extraction of entities and relationships, followed by review and verification by language experts to ensure knowledge accuracy.

3.2.2. Integrated scheduling engine

The workflow is as follows: After receiving user input, the system performs intent recognition. If the question falls under knowledge query categories, it converts the query into a structured query for the Knowledge Graph (KG) and initiates the RAG process. The system then retrieves relevant facts from the KG to construct an enhanced prompt, invokes the LLM inference service to generate answers, verifies factual consistency of the LLM output, and finally returns the final result to the user.

4. Implementation of the systems' core functional modules

Based on the aforementioned architecture, the system can achieve the following core teaching functions:

4.1. Personalized vocabulary learning path planning

This feature aims to address the questions of "what to learn" and "the learning sequence." Its implementation involves the following steps: First, the system evaluates the learners initial proficiency level or existing vocabulary and marks these as "known" nodes in the knowledge graph; next, it employs graph algorithms (such as Personalized PageRank) to identify the most relevant "unknown" vocabulary associated with these nodes, using criteria like root words, topics, or contextual similarity; finally, the LLM transforms this recommended vocabulary list into engaging and motivating learning tasks while generating personalized learning suggestions.

4.2. Contextualized and interactive vocabulary analysis

Traditional dictionary-based explanations are isolated and dull, whereas this system aims to deliver an immersive, conversational vocabulary learning experience. Its implementation works as follows: When a learner queries a word (e.g., "authentic"), the system initiates the RAG (Retrieval, Augmentation, Generation) process. First, it performs a retrieval from the knowledge graph to extract structured information related to "authentic," including part of speech (adjective), multiple definitions (① underground tunnel; ② authentic, standard), synonyms (authentic, genuine), antonyms (poor), common collocations (authentic Beijing dialect), typical example sentences, and relevant cultural context ("This term is frequently used when evaluating food or language"). Next, the information is augmented into a rich contextual framework. Finally, the LLM generates a natural, comprehensible, multi-perspective explanation based on this context, actively guiding the interaction.

4.3. Intelligent Q&A and confusion point analysis

Learners often encounter confusing nuances in vocabulary acquisition that are difficult to distinguish, such as "opportunity" versus "chance," or "suddenly" versus "sudden." For such differentiation tasks, the system employs the RAG strategy.

The knowledge graph provides core definitions, etymology, grammatical functions (e.g., "suddenly" primarily functions as an adverb, while "sudden" can serve as an adjective), and authoritative example sentences for each term. The LLM then translates and explains this structured comparative information using accessible language and apt analogies, generating comparative tables, scenario-based dialogues, and other forms of analytical content to help learners gain a deep understanding of the differences.

4.4. Dynamic practice and evaluation generation

Effective practice and assessment are crucial for consolidating learning outcomes. The system dynamically generates various types of exercises: multiple-choice questions based on the synonymy, antonymy, and hierarchical relationships of vocabulary in the knowledge graph; cloze questions utilizing lexical collocations; and reading comprehension or sentence correction tasks generated by the LLM, tailored to learners' common errors and incorporating specific vocabulary items. Upon completion of an assessment, the system not only assigns scores but also employs the LLM to analyze error causes and provides links to relevant knowledge points in the knowledge graph for reinforcement.

5. System evaluation methods and indicators

The evaluation of such a complex system must be multidimensional, encompassing both technical performance validation and effectiveness measurement in educational applications.

5.1. Technical performance evaluation

The basic model capability evaluation utilizes publicly available authoritative Chinese benchmark datasets such as C-Eval, CMMLU, and CLUE to assess the foundational abilities of the LLM in Chinese language comprehension and knowledge acquisition. Knowledge graph quality evaluation involves manually sampling and assessing the accuracy and recall rates of triples within the graph. The integrated question answering system evaluation constructs a question answering test set based on Chinese lexical knowledge (similar to the format of SQuAD-zh) to compare the accuracy, factual consistency, and anti-alience capabilities of both the "pure LLM" and the "KG-LLM integrated system" in answering questions.

5.2. Teaching effectiveness and user experience evaluation

This constitutes the core of evaluating the systems application value, requiring a rigorous experimental design. The experimental protocol involves recruiting a group of internationally comparable Chinese language learners and randomly assigning them to either an experimental group or a control group. The control group employs traditional learning methods (e.g., vocabulary memorization apps or online dictionaries) or a "pure LLM" teaching assistant without knowledge graph enhancement features. Over several weeks, vocabulary acquisition tasks are conducted, with learning outcomes assessed through pre-tests and post-tests.

5.2.1. Learning outcomes measurement indicators

The learning effectiveness evaluation indicators include: vocabulary acquisition volume and speed (by comparing the increase in vocabulary between the two groups before and after the experiment using vocabulary tests); vocabulary application ability (assessed through tasks such as contextual word completion, sentence construction, and short essay writing to evaluate learners mastery of vocabulary meanings and usage); and knowledge retention rate (evaluated through follow-up tests conducted sometime after the experiment to measure long-term memory retention).

5.2.2. User experience evaluation metrics

The user experience evaluation metrics include: user satisfaction and engagement, where feedback on system usability, enjoyment, and learning support is collected through standardized questionnaires (e.g., the System Usability Scale, SUS) and semi-structured interviews; Interaction Log Analysis, which systematically examines user behavior data on the platform, including average learning duration, feature usage frequency, problem types, and support requests, to quantify user engagement and identify strengths and weaknesses in system design; and statistical analysis employing methods such as t-tests or ANOVA to determine whether differences between the experimental and control groups across metrics are statistically significant ($p < 0.05$).

6. Challenges and future prospects

6.1. Challenges faced

First, constructing and maintaining high-quality knowledge graphs is a massive undertaking requiring sustained investment and interdisciplinary expertise. Second, regarding the depth and efficiency of integration algorithms, while current RAG models are effective, they still represent relatively superficial integration. Achieving deeper, bidirectional, and dynamic interaction and reasoning between LLMs and KGs remains a cutting-edge algorithmic challenge. Third, in terms of personalized model accuracy, precisely modeling learners' cognitive states, forgetting curves, and learning styles to deliver truly "tailored" adaptive instruction poses significant technical hurdles. Fourth, there is a lack of standardized evaluation frameworks; currently, no publicly available benchmarks or datasets specifically designed for such integrated teaching systems exist, making horizontal comparisons between different systems difficult.

6.2. Future outlook

First, multimodal integration: Incorporating diverse modalities, including text, images, audio (standard pronunciation), and short videos (demonstrating pragmatic scenarios), into knowledge graphs and instructional interactions to create a comprehensive immersive learning experience. Second, cognitive diagnosis and adaptive intervention: Leveraging cognitive science theories to develop refined learner diagnostic models that enable the system to proactively and in real-time intervene and correct misunderstandings or confusions regarding specific knowledge points. Third, affective computing and companion learning: Utilizing affective computing capabilities to monitor learners' emotional states (such as confusion, fatigue, or excitement) and dynamically adjust teaching strategies and interaction methods, thereby transforming the system into a more empathetic learning companion. Fourth, open-source ecosystem and community collaboration: Promoting the development of open-source Chinese vocabulary knowledge graphs, evaluation datasets, and system frameworks, encouraging global researchers and developers to participate collectively, accelerating technological iteration and practical implementation, ultimately benefiting international Chinese learners worldwide.

7. Conclusion

The integration of knowledge graphs and large language models offers an unprecedented technological opportunity to address fundamental challenges in international Chinese vocabulary instruction. The intelligent teaching system proposed in this paper combines the precision and structural integrity of knowledge graphs with the generative capabilities and interactivity of large language models, aiming to establish a more personalized, context-driven, and intelligent learning paradigm. Although challenges remain in graph construction, algorithmic integration, and evaluation frameworks, the continuous maturation of these technologies will undoubtedly exert a profound impact on international Chinese education and the broader field of intelligent education. Future research should focus on implementing concrete systems, empirically validating teaching effectiveness, and fostering collaborative development of open resources, thereby transforming this advanced concept into a powerful tool that genuinely empowers Chinese learners worldwide.

Funding

Phased achievement of 2025 Hainan Philosophy and Social Science Planning Project (Project No.: HNSK(ZC)25-238): A Study on Language Service Industry amid Hainan Free Trade Port Customs Closure and the University-Enterprise Cooperative School-running Model for International Chinese Language Education.

Disclosure statement

The author declares no conflict of interest.

References

[1] Cai W, Chu C, Cui X, et al., 2025, DeepSeek Empowering Innovation and Development in International Chinese Education. World Chinese Language Teaching, 2025(3): 291–315.

[2] Cao G, Liang Y, 2023, Construction and Application of a Knowledge Graph for International Chinese Education: A New Approach to Implement Large-Scale Tailored Teaching. Journal of Yunnan Normal University (Chinese as a Foreign Language Teaching and Research Edition), 2023(4): 5–15.

[3] Lian W, Xu J, 2025, Research on an Intelligent Q&A System for International Chinese Education Based on GraphRAG. Research on Modernization of Chinese Language Teaching – 2025 Special Collection (Volume II).

[4] Yu M, 2023, Research on the Development of Vocabulary Learning Resources for Adaptive Learning Systems. Research on Modernization of Chinese Language Teaching – 2023 Special Collection.