International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-14
© Cerebration Science Publishing
https://doi.org/10.70917/ijcisim-2026-1899
Article

An Empirical Study on the Effectiveness of an AI-Assisted Teaching Model for Chinese Constructional Grammar in Mobile Applications: Integrating Cognitive Load Theory in Second Language Acquisition

Jiayi Yang [1,*] and Xu Wang[2]

[1] School of Foreign Languages, North China University of Water Resources and Electric Power, Zhengzhou, Henan, 450046, China
[2] School of Date Science, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, China
* Correspondence author: Leonor1025bh@163.com

Abstract: Chinese construction grammar emphasizes the pairing of form and meaning and presents a teaching challenge in second language acquisition. This paper designs a teaching model for Chinese construction grammar based on the cognitive load theory of second language acquisition. First, a lightweight knowledge graph for Chinese construction grammar was constructed, defining four types of entities—construction names, concepts, examples, and structural components—along with their relationships, and mapping out the semantic network among constructions to reduce learners' internal cognitive load. By embedding the knowledge graph into a learning platform and linking knowledge points, the model accurately identifies students' knowledge gaps and recommends related concepts and practice exercises. By incorporating cognitive load theory in second language acquisition, we completed the construction of the Chinese construction grammar teaching model. Performance test results show that, compared to the best-performing KGAT model, the recommendation algorithm proposed in this paper achieved a 0.26% and 0.03% increase in recall, respectively, and a 0.18% and 0.35% increase in NDCG, respectively. Empirical experimental results show that the experimental group demonstrated significant improvements in both receptive and productive skills, while the control group showed a significant improvement only in receptive skills. This demonstrates that the application of the AI-assisted Chinese constructional grammar teaching model on mobile devices effectively reduces students' cognitive load and enhances their learning efficacy in Chinese constructional grammar.

Keywords: Chinese constructional grammar; knowledge graph construction; exercise recommendation; cognitive load in second language acquisition

1. Introduction

With the deepening of globalization, Chinese, as one of the key languages for international communication, has seen its teaching and research receive increasing attention [1–2]. Construction grammar emerged in the 1980s; it emphasizes the pairing of form and meaning and advocates a construction-centered approach, viewing constructions as wholes whose overall function exceeds the sum of their parts, rather than as a simple aggregation of the meanings of their individual components [3–6]. The rise of construction grammar has provided a fresh perspective on language acquisition and teaching in Chinese as a second language, breaking away from the traditional "verb-centered" grammatical paradigm and revealing the uniqueness and complexity of Chinese grammar [7–10].

However, traditional construct-based instruction fails to achieve the dynamic emergence and high-frequency reinforcement of constructs in complex contexts. With the development of mobile devices, mobile learning has demonstrated significant advantages in terms of convenience, personalization, and richness in accessing educational information, resources, and services, thereby providing support for addressing the aforementioned issues [11–14]. Nevertheless, most mobile applications remain at the rule-based level. Against this backdrop, empirical research on the effectiveness of AI-assisted Chinese construction grammar teaching models in mobile applications holds significant practical significance.

This study proposes an AI-assisted teaching model for Chinese construction grammar. First, a knowledge graph learning platform was developed to clearly present the relationships between knowledge points in the form of a semantic network, thereby reducing the learners' cognitive load. Next, students' knowledge levels were assessed based on their basic information, learning objectives, and specific learning progress. Exercises are recommended based on the relationships among knowledge points, the relationships between knowledge points and exercises, and the user's mastery of the knowledge points, thereby improving the accuracy of exercise recommendations. Based on cognitive load theory in second language acquisition and integrating AI technologies such as knowledge graphs and recommendation algorithms, a teaching model for Chinese construction grammar was constructed. A comparative experiment was designed to validate the algorithm's performance, and 100 second-year students majoring in non-Chinese disciplines at a university in Province A were selected as research subjects to conduct an empirical validation of the effectiveness of this teaching model.

2. Building a Knowledge Graph for Chinese Constructional Syntax

2.1. Constituent Syntax

The core of construction grammar lies in viewing "constructions" as the basic units of language. A construction is a unity of form and meaning, and its overall meaning is not simply the sum of its parts. The main characteristics of constructions include:

Holism: The meaning of a construction is not the simple sum of its parts, but rather a specific semantic meaning that emerges from the whole. For example, "going shopping" is not a literal combination of "going" and "shopping," but rather conveys the complete semantic meaning of "strolling along the street, browsing store windows, shopping, or relaxing."

Generativity: A structural framework can generate new linguistic expressions through expansion, analogy, and other means. For example, the "有 V 有 V" (a Chinese Phrases) construction can be expanded to include different verbs. This construction is used to express the simultaneous occurrence of multiple related actions or states, creating a rich and varied scenario.

Inheritance: The meaning and usage of constructions are often inherited from other related constructions. For example, there is a clear relationship of semantic and usage inheritance between "V 得 C" (a Chinese Phrases) and "V 得怎么样" (a Chinese Phrases; both evolve from describing a specific state to forming a general inquiry, illustrating the inheritable nature of constructions.

2.2. Building the Overall Process

Drawing on the experience of previous researchers and taking into account the characteristics of a Chinese construction grammar knowledge graph, this paper proposes a five-step method for constructing a Chinese construction grammar knowledge graph, namely: necessity analysis, acquisition of Chinese construction grammar data, knowledge graph design, knowledge storage, and knowledge visualization. The construction process for the Chinese construction grammar knowledge graph is shown in Figure 1. A brief explanation follows:

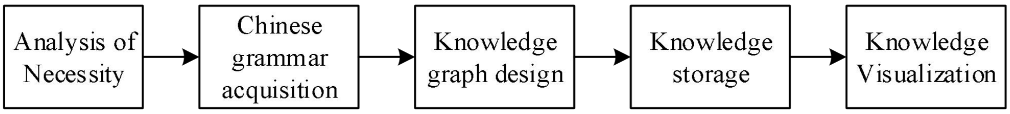

Figure 1. Flow chart of constructing Chinese grammar knowledge graph

2.3. Acquisition of Chinese Constructional Syntax Data

There are two primary sources of data on Chinese construction grammar: authoritative grammar textbooks and Baidu Baike. Grammar textbooks are the most direct way for users to obtain grammatical data, and they represent the essence of knowledge refined through continuous research and refinement by experts. Since the knowledge points in Chinese construction grammar are relatively fixed compared to other fields, this paper primarily obtains grammatical data from textbooks written by leading experts in various areas of Chinese construction grammar.

Manually extracting definitions and examples of parts of speech directly from text-based books is too labor-intensive; therefore, this paper also selected Baidu Baike—a website rich in knowledge—as a source for relevant Chinese construction grammar corpora. The primary focus is on obtaining definitions and examples of parts of speech.

2.4. Knowledge Graph Design

The design of a knowledge graph is equivalent to the design of the conceptual layer, which includes entity design and relationship design. When designing a knowledge graph, it is essential to ensure the completeness and accuracy of entities, strive to design a lightweight knowledge graph, and filter out irrelevant data.

Based on actual business scenarios and the Chinese construction syntax data already available in the local database, this paper categorizes entities into four types: construction names, construction concepts, construction examples, and construction structural components.

Based on these four entity types, this paper manually designed corresponding entity relationships, which are divided into the following four major categories:

1) “A contains B” relationship

The “A contains B” relationship indicates that entity A contains entity B; it primarily describes the relationship between different construct names.

2) “A is a concept of B” relationship

The “A is a concept of B” relationship indicates that entity A is a concept of entity B; it primarily establishes a correspondence between construct concepts.

3) “A is an instance of B” relationship

The “A is an instance of B” relationship indicates that entity B is an instance of entity A; it primarily establishes a correspondence between a construct name and its specific instances.

4) “Structure A is associated with B” relationship

The “Structure A is associated with B” relationship indicates that a certain association exists between entity A and entity B; it is primarily used to describe the semantic relationship between the various components within a construction.

2.5. Knowledge Repository

After completing the conceptual model design for the knowledge graph, the data stored in the local database must be populated into the conceptual model for knowledge storage. Knowledge storage consists of two parts: first, standardizing the data format; and second, importing the data into the Neo4j graph database.

Since the format of the Chinese construction syntax data initially obtained still differs somewhat from the data format required by the knowledge graph’s conceptual model, it is necessary to standardize the data format. Standardizing the data format involves unifying both the entity data format and the relationship data format.

In recent years, graph databases have become the predominant choice for knowledge graph storage. The Neo4j graph database has gained widespread adoption due to its flexible design, high performance, and ease of use. In the Neo4j graph database, nodes represent entities, and edges represent relationships between entities. To start the Neo4j service, enter “Neo4j.batconsole” at the command prompt; the port is 7474. Once the URL with port 7474 appears, you can access the Neo4j graph database.

When importing data, you can use the official Neo4j-import tool for batch imports, or use the LOAD command in the Neo4j graph data web interface. Regardless of which method you choose, you must save each type of entity and relationship as a CSV file before importing.

Once you have the correct CSV files, place them in the `import` folder within the Neo4j installation root directory to facilitate the subsequent import. When importing files in bulk, issues with Chinese characters may arise, preventing them from displaying correctly in the graph. Since the names of knowledge points in the Chinese Constructive Syntax Knowledge Graph consist entirely of Chinese characters, you must adjust the encoding format of the CSV files.

This article uses the LOAD command in the Neo4j graph database to import CSV files in bulk. When using the LOAD command to import Neo4j graph data, you must also specify the names of the relationships.

2.6. Knowledge Visualization

Once a CSV file has been successfully imported into the Neo4j graph database, Neo4j displays a limited amount of knowledge graph data via a graphical interface.

To view detailed information about the graph, you can directly enter a Cypher query to explore specific construct entities and relationships.

The study of Chinese construction grammar emphasizes syntactic awareness and grammatical structure. Building a knowledge graph for Chinese construction grammar essentially presents users with the network relationships among Chinese construction concepts, helping to clarify the formal connections and distinctions between different constructions; at the same time, the graphical interface helps enhance learners' motivation and attention. Furthermore, the knowledge graph designed in this chapter will serve as the data foundation for Chinese construction grammar teaching models, laying a solid groundwork for the design of such teaching approaches.

3. Design of Intelligent Recommendation Algorithms

The intelligent recommendation features in traditional Chinese language learning systems have the following issues: First, they cannot accurately assess a user's mastery of specific knowledge points. Most systems simply identify questions answered incorrectly during practice as unmastered knowledge points; however, this method does not fully account for all possible scenarios. Therefore, this system evaluates a user's level of understanding of knowledge points by analyzing their basic information and initial learning goals, combined with their learning attitude and performance during the learning process; Second, existing systems do not recommend exercises based on the relationships between knowledge points, the relationships between knowledge points and exercises, or the mastery levels of the same knowledge point among different users. This is because knowledge points in Chinese language learning are interconnected; a user's failure to master one knowledge point may be caused by a lack of mastery of another. Additionally, similar users may share similar knowledge gaps. Therefore, incorporating these factors into exercise recommendations would yield relatively ideal results. The entire recommendation model is shown in Figure 2.

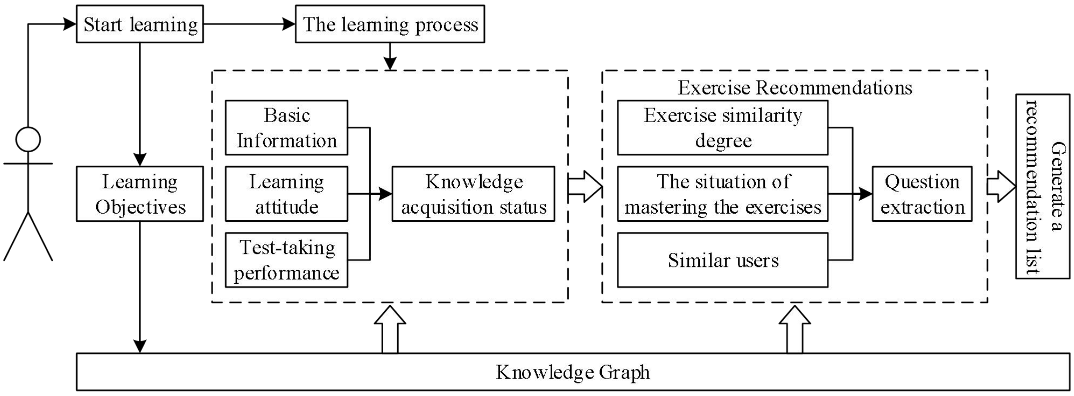

Figure 2. Recommend the overall process

3.1. Calculation of Knowledge Mastery

Users' learning data within the system forms the foundation for intelligent recommendations. By analyzing this answer data, the system can determine users' mastery of specific knowledge points and identify their areas of weakness. By combining users' incorrect answer data with the generated Chinese knowledge graph, we can assess their mastery of specific knowledge points. This mastery reflects students' cognitive levels. We determine a user's mastery of a knowledge point based on their test performance, where answering a question correctly is considered a certain level of mastery of that knowledge point. Correct answers are categorized as either "guessed correctly" or "answered correctly"; to accurately calculate the level of mastery, "guessed correctly" cases must be excluded from the calculation. Let $Q=\{Q_1,Q_2,\ldots,Q_n\}$ denote the set of exercises, $m_u=\{m_{u1},m_{u2},\ldots,m_{un}\}$ the vector representing mastery of knowledge points, and q the knowledge point matrix. Then, the status of unanswered exercises is given by the following equation:

$$\rho_{uv} = \prod_{N=1}^{N} m_{un}^{q_{vn}} = f\left(m_u, q_v\right) \tag{1}$$

The function $f\left(m_u, q_v\right)$ represents the power operation for m and q, and q_{vn} takes the values 0 or 1. The following expression represents the probability that the user has guessed the answer to the exercise:

$$P\left(R_{uv} = 1 \mid \rho_{uv} = 0\right) = G_v \tag{2}$$

In the formula, P represents the student, $R_{uv} = 1$ represents the set of questions the student answered correctly, $R_{uv} = 0$ represents the set of questions the student answered incorrectly, and G_v represents the probability that the student guessed the answer to question Q_n. In addition to the guessing probability, it is also necessary to classify questions by difficulty level; the higher a user's pass rate for a question, the lower its difficulty. The calculation formula is as follows:

$$A_i = 1 - \frac{\overline{X_i}}{S_i} \tag{3}$$

In the formula above, A_i represents the pass rate for the question, $\overline{X}_i$ represents the average score for the question, and S_i represents the point value of the question. A higher pass rate indicates that the question is easier. In this paper, a question with a pass rate less than 0.6 is defined as difficult, a question with a pass rate less than 0.8 is defined as of medium difficulty, and a question with a pass rate of 0.8 or higher is defined as easy. Once the difficulty levels are defined, the following formula can be used to calculate a user's mastery of a specific knowledge point:

$$E(U,K) = \alpha R_m + \beta R_n + \gamma R_s - G_v \tag{4}$$

In the above formula, $E(U,K)$ represents student U 's mastery of knowledge point K, where α represents the weight of easy-level questions, β represents the weight of medium-level questions, and γ represents the weight of difficult-level questions, while R also represents the correct answer rate.

3.2. Calculating the Similarity Between Test Items

This system calculates the similarity between a user's incorrect answers on a given knowledge point and recommended questions on the same knowledge point in the knowledge graph to identify the most relevant practice questions, which are then recommended to the user. This approach enables the system to recommend practice questions based on the user's knowledge gaps. The formula for calculating question similarity is shown below:

$$W_{AB} = \frac{\sum_{i=1}^{n} \left(X_i * Y_i\right)}{\sqrt{\sum_{i=1}^{n} \left(X_i\right)^2} * \sqrt{\sum_{i=1}^{n} \left(Y_i\right)^2}} \tag{5}$$

In the above formula, X_i and Y_i represent the vectors of two practice problems, and W_{AB} is the similarity between problems *X* and *Y*. By calculating the similarity between them, we can identify practice problems similar to the incorrect one.

3.3. Combine Similar User-Generated Recommendation Lists

Based on users' answering patterns in the learning system, users who have made the same mistakes may be more interested in questions of the same type; that is, similar users are more likely to have similar knowledge gaps. This process primarily involves two steps: first, identifying a set of users whose areas of interest are similar to those of the target user; then, recommending exercises from this set that the target user has not yet encountered but that users in the set find interesting. Exercise recommendations are made by calculating the similarity between users. The formula for calculating user similarity is as follows:

$$W_{ij} = \frac{|N(i) \cap N(j)|}{\sqrt{|N(i)||N(j)|}} \tag{6}$$

Here, $N(i)$ represents the set of incorrect answers for user i, and $N(j)$ represents the set of incorrect answers for user j, So, $|N(i) \cap N(j)|$ represents the set of questions that both i, j got wrong, $|N(i)N(j)|$ represents the total number of questions that i, j got wrong, and W_{ij} represents the similarity between user *i* and user *j*. Once the user similarity is obtained, the next step is to calculate the degree of interest in the questions that users got wrong. The following formula can be used for this calculation:

$$P(u,i) = \sum_{v \in S(u,K) \cap N(i)} W_{uv} r_{vi} \tag{7}$$

In particular, $S(u,K)$ consists of the K users whose interests are most similar to those of user u, while $N(i)$ is the set of users who have interacted with exercise i, W_{uv} is the similarity in interests between user *u* and user *v*, and r_{vi} represents user *v*'s interest in the practice questions. Since we are using implicit feedback data based on a single behavior, all $r_{vi} = 1$ values can be simplified using the following formula:

$$p(A,e) = W_{AB} * R_{B_e} + W_{AC} * R_{C_e} + W_{AD} * R_{D_e} \tag{8}$$

Here, $P(A,e)$ represents A's interest in e, W_{AB} represents the similarity between A and B, R_B represents B's interest in e, and so on. Using these values, we can calculate a list of recommended questions for similar users.

3.4. Design of a Question-Selection Algorithm

Before generating intelligent recommendations, practice questions must first be selected. This paper uses a knowledge graph to select questions, which must be based on key concepts, question difficulty, and the user's answer history. The selected questions must not be duplicates, and the question types must be evenly distributed. Heuristic random search algorithms use heuristic information about the problem to guide the search process, thereby narrowing the search scope and reducing the problem's complexity. By evaluating each search location to find the optimal one, these algorithms can eliminate a large number of ineffective search paths and improve search efficiency. When selecting practice questions, they are grouped based on factors such as the difficulty level of the knowledge points and question types, and questions are sampled in varying proportions. The system then uses the question information to identify and select those that meet the requirements. Figure 3 illustrates the question selection process:

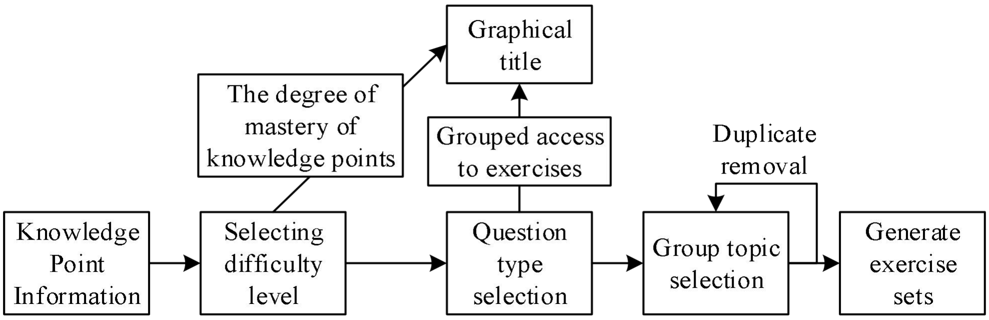

Figure 3. Flow chart of drawing questions

3.5. Smart Recommendation Design

Using knowledge graphs to provide intelligent recommendations for concepts and practice exercises can effectively meet students' learning needs. When a student fails to master a particular concept during the learning process, they can look up related concepts and exercises in the knowledge graph. For example, if a user has not mastered a particular verb tense while learning Chinese, it may be because they have not thoroughly understood the relevant morphological or syntactic concepts. Therefore, the system can recommend related concepts and exercises to guide the user through targeted practice, preventing them from spending excessive time on ineffective study. The overall process of intelligent recommendation is shown in Figure 4.

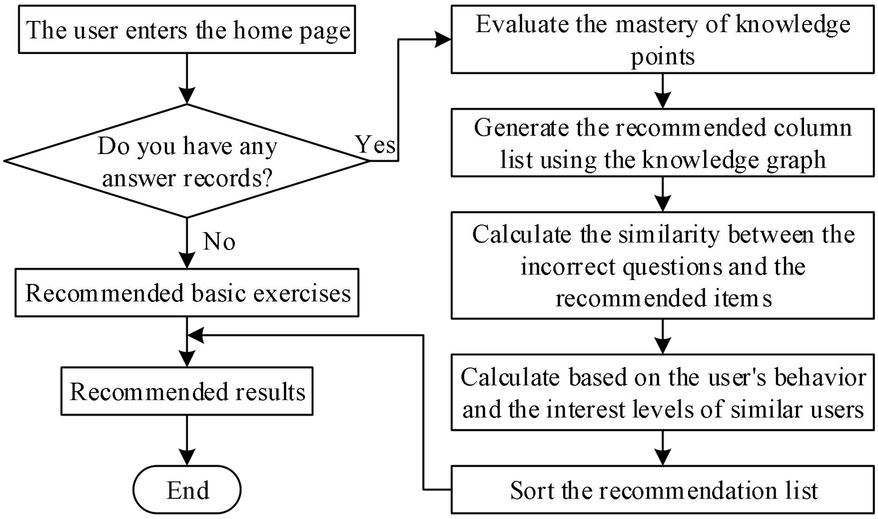

Figure 4. Inteligent recommendation process

3.6. Recommended Performance Tests

This paper uses two publicly available datasets to demonstrate the superiority of the improved model:

1) CG-KG is a dataset specifically designed for recommending Chinese construction grammar questions, provided by the Chinese Language Resources Alliance. It contains attribute information such as question stems, answer choices, correct answers, and associated grammar points.

2) HSK-Grammar is a dataset containing records of interactions between Chinese learners and grammar questions, primarily collected from the HSK (Chinese Proficiency Test) practice platform. The datasets record learners' historical answer sequences, with each question annotated with the tested grammatical point, difficulty level, and question type.

This paper employs three widely used evaluation metrics to assess the effectiveness of Top-K recommendations and preference ranking: recall, normalized discounted cumulative gain (NDCG), and precision.

To demonstrate the effectiveness of the proposed method, the KGAT-IF model is compared with the following baseline models. Experiments were conducted in a Python 3.6 environment. The models are described as follows:

1) FM is a benchmark factorization model that learns second-order feature interactions in a linear manner.

2) NFM is an advanced factorization model that incorporates FM into a neural network.

3) CKE integrates auxiliary information—such as structural information, text data, and image data—into the recommendation system. It extracts the structural representation of items via TransR to improve the quality of the recommendation system.

4) The CFKG model utilizes heterogeneous entity embeddings for recommendations and proposes a soft matching algorithm based on an embedding knowledge base to generate personalized explanations for recommended items.

5) MCRec performs Top-N recommendations using a meta-path-based context and neural co-attention model, which extracts eligible meta-paths as connections between users and items.

6) The RippleNet model updates user embeddings using entity node embeddings and iteratively expands users' interests along the knowledge graph through preference propagation to discover their latent preferences.

7) GC-MC uses a GCN encoder to encode graph-structured data.

8) The KGAT model combines user interaction history with a knowledge graph to model higher-order relationships in the graph in an end-to-end manner; it uses attention mechanisms to distinguish the importance of neighboring nodes and optimize the propagation of embeddings between nodes.

The recommendation results of the proposed model and the above models were evaluated on two datasets using the two metrics of recall and NDCG; the comparison results are shown in Table 1.

The table reflects the recall and NDCG values of each model on the two datasets. As shown in the table, the model proposed in this paper outperforms the comparison models on both datasets. Specifically, on the CG-KG and HSK-Grammar datasets, compared to the best-performing comparison model (KGAT), the proposed model achieves a 0.26% and 0.03% increase in recall, respectively, and a 0.18% and 0.35% increase in NDCG, respectively.

Table 1. xperimental comparison

Model	**CG-KG**		**HSK-Grammar**	
	Recall	**NDCG**	**Recall**	**NDCG**
FM	0.1341	0.0891	0.0775	0.118
NFM	0.1371	0.092	0.0826	0.1212
CKE	0.1344	0.0881	0.0734	0.1191
CFKG	0.1141	0.0776	0.0720	0.1143
MCRec	0.1113	0.0781	--	--
RippleNet	0.1331	0.0910	0.0793	0.1240
GC-MC	0.1318	0.0881	0.0820	0.1252
KGAT	0.1488	0.1004	0.0874	0.1320
This model	0.1514	0.1022	0.0877	0.1355

The results of personalized exercise recommendations are shown in Table 2, which tests the recommendation accuracy for similarity coefficients of 0.5, 0.7, and 0.9, respectively. When the similarity coefficient is 0.5, the accuracy outperforms that achieved with coefficients of 0.7 and 0.9. This is because a lower similarity coefficient allows the recommendation algorithm to retrieve more questions from the knowledge graph that are semantically related to the target grammar point but not identical, thereby covering grammar points that learners may not yet have mastered but that still hold recommendation value, thus improving the accuracy of the recommendations.

Table 2. Personalized review recommendation

		Accuracy	
Recommended number	0.5	0.7	0.9
5	0.8588	0.8098	0.7577
10	0.8590	0.8083	0.7604
15	0.8671	0.8174	0.7683
20	0.8770	0.8234	0.7719
25	0.8891	0.8385	0.7885
30	0.8959	0.8449	0.7951
35	0.8994	0.8516	0.8014
40	0.9126	0.8652	0.8136
45	0.9258	0.8757	0.8257
50	0.9339	0.8837	0.8315

4. Developing a Knowledge Graph-Based Teaching Model for Chinese Construction Grammar

4.1. Knowledge Graph Learning Platform

The knowledge graph learning platform selected for this study is an online learning platform based on various artificial intelligence technologies, including knowledge visualization, natural language processing, and personalized recommendation algorithms. It integrates a Chinese language subject knowledge graph with a question bank organized around the knowledge graph. The platform includes two user interfaces—one for teachers and one for students—which provide support for both teaching and learning. It offers a variety of functions, including knowledge point retrieval, knowledge relationship queries, knowledge graph visualization, question selection and test generation, practice exercises from the question bank, and learning progress analysis, thereby supporting the entire teaching process.

1) Knowledge Point Retrieval

Knowledge point retrieval is the process whereby a user enters a natural language query regarding a specific Chinese construction grammar knowledge point, and the learning platform provides the answer to that query. Users enter the knowledge query they wish to search for in the "Search Box" and then click "Search." The system retrieves relevant information linked to the target knowledge point in the background and displays it in the front-end search results bar, accurately returning the search results to the user. If a user enters a question type that the system cannot recognize or a knowledge point not yet included in the system, the system will display an error message.

2) Knowledge Relationship Query

A knowledge relationship query refers to the process in which, after a user enters a knowledge entity to be queried, the platform visually displays the relationships between that entity and other related knowledge entities, presenting the list of relationships in the form of triples. In the knowledge relationship query module, the knowledge graph learning platform combines graphical representations to provide users with two presentation formats for query results: "knowledge relationship diagrams" and "relationship lists." This allows users to clearly and intuitively identify other knowledge entities associated with the target knowledge entity, as well as the relationships between them.

3) Knowledge Graph Visualization

Building on specific knowledge relationship queries, to help users better grasp the relationship between the whole and parts of the knowledge graph, the learning platform also provides a feature for visualizing the full picture of the knowledge graph. Under the Knowledge Graph Visualization navigation bar, clicking the "Full Knowledge Graph View" button causes the system to visually display all knowledge entities and relationships contained in the global knowledge graph within the "Graph Display Panel." "Circular entities" of different colors represent different entity levels, while "linear relationships" of different colors denote different relationship categories.

4) Exercise Resource Library

Building on the knowledge point and relationship query functions, the exercise resource library is another core feature of the knowledge graph learning platform. To meet the distinct needs of teachers and students, the platform has designed different functional modules for each user interface: the "Select Questions and Create Tests" feature is available on the teacher side, while the "Practice from Question Bank" feature is available on the student side. In the "Question Bank Practice" module, the platform provides students with personalized question recommendations. By analyzing data generated during the user's learning process and integrating it with the knowledge graph, the platform assesses the user's mastery of specific knowledge points, thereby providing intelligent recommendations based on areas of weakness and suggesting appropriate learning content and practice questions. The "Question Selection and Test Generation" module provides teachers with a question-selection strategy based on incorrect answers. By analyzing the records of incorrect answers for the entire class or individual students, the platform uses personalized recommendation algorithms and other technologies to intelligently recommend questions and automatically generate test papers. Teachers can then manually refine these recommendations by adding or removing questions from the automatically generated worksheets based on their experience and judgment, ultimately creating formal tests or homework assignments. These can be distributed to students by clicking "Publish."

5) Learning Progress Analysis

Knowledge graphs not only link knowledge points and learning resources to enable the structured organization of resources, but also systematically collect user behavior data and map it to resource objects, thereby facilitating precise analysis of learning progress. By extracting and analyzing user platform usage data, knowledge graph-based learning platforms can effectively enhance user descriptions and representations, build learner models, accurately diagnose learning progress, and consequently improve the accuracy, diversity, and interpretability of resource recommendations.

6) Other Features

To better support teachers and students in managing assignments, in addition to the features mentioned above, a "Homework Center" module has been added to the teacher interface, while the student interface now includes "My Assignments" and "Mistakes Notebook" modules.

4.2. Developing a Teaching Model for Chinese Constructional Syntax

To develop an effective model for teaching Chinese construction grammar that is both universally applicable and provides practical guidance, this study selected the cognitive load theory of second language acquisition for an overview and analysis.

Cognitive load theory analyzes the design and effectiveness of cognitive activities from the perspective of cognitive resource allocation. Most existing studies attempt to regulate task-induced cognitive load within the cognitive load framework by varying instructional methods, thereby effectively mobilizing learners' limited cognitive resources to optimize learning tasks and enhance learning outcomes. Recent research has primarily focused on the field of multimedia-assisted instruction. Based on practical teaching considerations, attempts have been made to reduce external cognitive load by altering the organizational methods of material presentation, with the aim of optimizing learning outcomes; reducing both internal and external cognitive load may promote learning. Research on cognitive load in the field of second language acquisition is still in its early stages; however, because it may provide concrete guidance for instructional design, it has become an emerging research topic, with relevant studies in both the input and output phases of language learning.

Compared to the first language, second language processing places higher demands on cognitive resources. By combining subjective evaluations with electroencephalogram (EEG) measurements to assess cognitive load, a study compared the cognitive load of native speakers and foreign language learners while reading Korean texts. The findings revealed that native speakers experienced lower cognitive load than foreign language learners, and that the cognitive load of foreign language learners increased as text difficulty increased. Most existing discussions on related topics in the field of second language acquisition have followed the tradition of cognitive load research, exploring the role of strategies such as schema activation, new word annotation, and visual aids during listening in facilitating text comprehension among foreign language learners. Essentially, these strategies aim to enhance text comprehension by reducing the cognitive load involved in text processing for second language learners.

Compared to research on the input phase, there has been relatively little research on cognitive load in spoken and written output. Although existing studies have primarily focused on output tasks, their areas of interest are relatively scattered, ranging from how to reduce cognitive load through temporal regulation to improve output quality to research on the cognitive demands generated by task design. Although research on cognitive load in the output phase is still in its infancy, it is closely related to the classic topic of task-based teaching in the field of second language acquisition.

The teaching of Chinese constructional grammar supported by a knowledge graph learning platform is, in fact, an effective application of blended learning. This paper prioritizes the design of instructional objectives and, based on these, plans instructional activities and resources accordingly. During in-class instruction, guided by the principles of cognitive load theory regarding the control of information presentation and task complexity—and leveraging the functionalities of the knowledge graph learning platform—students engage in either self-directed inquiry activities or group-based collaborative inquiry activities tailored to knowledge points of varying difficulty. This approach reduces unnecessary intrinsic cognitive load while enhancing relevant cognitive load.

In terms of instructional assessment, in addition to formative and summative assessments, an evaluation of the teaching implementation process has been incorporated. Furthermore, in line with the characteristics of this model, an evaluation of instructional resources has been added. In summary, the Chinese constructional grammar teaching model fully integrates the core concepts of cognitive load theory in second language acquisition.

5. An Empirical Study on the Effectiveness of Mobile Apps for Teaching Chinese Constructional Grammar

5.1. Study Population

This study involved 100 second-year students from two classes (Class A and Class B) in a non-Chinese language major at a university in Province A, with 50 students in each class. The average ages of the two classes were similar (20.2 and 20.3, respectively), and the proportion of female students was also roughly the same (70% and 72%). Preliminary tests indicated no significant differences between the two classes in terms of receptivity to and production of Chinese constructional grammar. During the experiment, students in the experimental group all used a mobile learning platform to complete their Chinese constructional grammar learning tasks.

5.2. Research Tools

The research tool for this study is a grammatical proficiency test. The test assesses grammatical structures in Chinese construction grammar. The test is divided into receptive and productive proficiency sections. The receptive proficiency section consists of multiple-choice questions. Each question presents a Chinese sentence containing the target grammatical structure, with the structure underlined. The correct answer is the correct name for that structure, while the distractors are names for similar structures. Each question is worth 1 point; correct answers earn 1 point, and incorrect answers do not result in point deductions. There are 30 questions in total. All tests are delivered and completed via a mobile app.

5.3. Data Collection and Analysis

This study spanned eleven instructional weeks. On the Monday morning of the first instructional week, the instructor administered a 90-minute grammar proficiency test as a pre-test. Subsequently, from the second through the tenth instructional week—a total of nine weeks—the teaching experiment was conducted, with the instructional objective focused on Chinese construction-based linguistic expressions. The control group followed a traditional teaching model, while the experimental group adopted an AI-assisted teaching model for Chinese construction grammar.

The control group received instruction on a new target grammar point every Monday for 45 minutes. After class, students were assigned an online grammar assignment to be completed within 30 minutes by midnight on Thursday; the assignment consisted of multiple-choice questions. Friday was reserved for a 45-minute review session, during which the instructor provided feedback and explained any issues encountered in the grammar assignment.

The experimental group was assigned a target grammar learning task every Monday, which students were required to complete online by midnight on Thursday. All online learning tasks had to be completed within 75 minutes and could be done in one sitting or in multiple sessions. An in-person grammar class was held every Friday, lasting 45 minutes.

On Monday of Week 11, a post-test is administered. The post-test covers exactly the same topics and question types as the pre-test; the question stems remain unchanged, but the order of the answer choices is rearranged. The time limit remains 90 minutes. After students submit their test answers via the mobile app, research team members grade both the pre-test and post-test papers, enter the scores for both tests into an Excel spreadsheet, and finally conduct quantitative analysis using the SPSSAU online data analysis platform.

5.4. Experimental Results

Table 3 presents the results of the independent samples t-test for the scores of the two classes on the receptive ability pre-test and the productive ability pre-test. Since the scores of the two classes on both the receptive pre-test (P = 0.092 > 0.05) and the productive pre-test (P = 0.746 > 0.05) did not show statistical significance, this indicates that Classes A and B scored consistently on both types of pre-tests, with no differences observed. It can therefore be concluded that there was no difference in the students' proficiency in Chinese constructional grammar between the two classes at the time of the pre-tests. Class A was then randomly assigned as the control group, and Class B as the experimental group.

Table 3. Test of acceptance and output performance

Name	Group (M±Std)		t	P
	A(N=50)	B(N=50)		
Acceptance before acceptance	14.72±2.01	13.95±1.78	1.712	0.092
Preoutput measurement	7.56±1.13	7.39±1.54	0.335	0.746

Paired-sample t-tests were conducted on the pre- and post-test scores for receptivity in both groups. The results are shown in Table 4.

The difference between the pre-test and post-test receptivity scores in the control group was statistically significant at the 0.01 level (P = 0.000). Specifically, the mean post-test receptivity score (15.63) in the control group was significantly higher than its mean pre-test score (14.72). The difference between the pre-test and post-test receptivity scores in the experimental group was also statistically significant at the 0.01 level (P = 0.000), and the mean post-test receptivity score (18.22) was significantly higher than the mean pre-test receptivity score (13.95).

Table 4. Test the test of the performance force and the posterior test

	Group (M±Std)		Difference value	t	P
	Pretest	Post-test			
The control group was tested before the matching	14.72±2.01	15.63±1.77	-0.91	-3.836	0.000**
The experimental group was tested before the match was tested	13.95±1.78	18.22±1.82	-4.27	-12.421	0.000**

Subsequently, *Cohen's d*—a measure of effect size—was calculated for both sets of data to compare the magnitude of score differences between the two classes on the pre- and post-tests of receptive ability. The results of the effect size analysis for the control and experimental groups are shown in

Table 5.

The *Cohen'sd* value for the difference between the pre- and post-tests in the control group (0.605) is greater than 0.5 and less than 0.8, indicating a moderate difference between the pre- and post-test data for the control group. The *Cohen'sd* value for the difference between the pre- and post-tests in the experimental group (1.961) is greater than 0.8, indicating a large difference between the pre- and post-test data for the experimental group. The results indicate that the post-test scores for receptivity in both groups were significantly higher than the pre-test scores; however, the magnitude of the difference between pre- and post-test scores in the experimental group was greater than that in the control group. It can be concluded that both the AI-assisted Chinese construct grammar teaching model and the traditional teaching model have significant effects on improving students' receptivity to Chinese construct grammar, but the AI-assisted model is more effective.

Table 5. The effect of the performance of the performance force

	Effect indicator				
Name	**Mean difference**	**Difference value 95% CI**	**df**	**Deviation deviation**	***Cohen'sd***
The control group was tested before the matching	-0.91	-1.456~-0.445	49	1.566	0.605
The experimental group was tested before the match was tested	-4.27	-4.882~-3.511	49	2.138	1.961

Paired-sample t-tests were conducted on the pre-test and post-test scores for productive ability in both groups. The results of the paired-sample t-tests for the pre-test and post-test scores of productive ability in the control and experimental groups are shown in Table 6. No significant difference was found between the pre-test and post-test scores for productive ability in the control group (p > 0.05). In the experimental group, a significant difference at the 0.01 level was observed between the pre-test and post-test scores for productive ability (p = 0.000), and the mean post-test score (9.34) was higher than the mean pre-test score (7.39).

Table 6. Test of output performance force before and after test

Name	**Group (M±Std)**		**Difference value**	**t**	**P**
	Pair 1	**Pair 2**			
The control group outputs the previous test	7.56±1.13	7.81±1.08	-0.25	-1.596	0.118
Experimental experimental test	7.39±1.54	9.34±1.48	-1.95	-10.699	0.000**

To conduct an in-depth analysis of the magnitude of the difference between the pre-test and post-test scores for productive skills in the experimental group, we calculated *Cohen'sd*—the effect size measure for the difference in pre- and post-test scores for productive skills in the experimental group. The effect size measures for the pre- and post-test scores of productive skills in the experimental group are shown in Table 7.

The effect size value (1.693) for the difference between the pre- and post-test scores of productive skills in the experimental group is greater than 0.8, indicating a significant difference in the pre- and post-test scores for productive skills in the experimental group. This indicates that there was no significant difference between the pre-test and post-test scores for the control group's productive ability. In contrast, the experimental group showed a significant difference between the pre-test and post-test scores for productive ability, and the magnitude of this difference was substantial. It can be concluded that the traditional teaching model has no significant effect on improving productive ability in Chinese grammar, whereas the AI-assisted teaching model for Chinese construction grammar has an extremely significant effect on enhancing productive ability in Chinese construction grammar.

Table 7. Experimental group output performance index

	Effect indicator				
Name	**Mean difference**	**Difference value 95% CI**	**df**	**Deviation deviation**	***Cohen'sd***
Experimental pretest test-Experimental post-test	-0.91	-2.316~-1.583	49	1.152	1.693

6. Conclusion

This study designed a teaching model based on knowledge graphs, personalized recommendation algorithms, and the cognitive load theory of second language acquisition. Through empirical research, the study evaluated the effectiveness of this model in enhancing learners’ ability to acquire and produce constructional grammar.

Comparative experimental results show that the model proposed in this paper outperforms existing models in terms of Recall and NDCG metrics on both the CG-KG and HSK-Grammar datasets. When compared with the best-performing model among the comparison models, the KGAT model, the proposed model achieved Recall improvements of 0.26% and 0.03%, respectively, while the corresponding NDCG improvements were 0.18% and 0.35%. The model achieves the highest recommendation accuracy when the similarity coefficient is 0.5; when the number of recommended questions is 50, the accuracy reaches 0.9339.

A teaching experiment compared the effectiveness of an AI-assisted Chinese construction grammar teaching model and a traditional teaching model in enhancing learners’ receptive and productive Chinese grammar skills. The experimental results show that both the traditional teaching model and the AI-assisted teaching model for Chinese construction grammar had significant effects on improving students’ receptive skills in Chinese grammar. The post-test results for both receptive and productive skills in the experimental group showed significant improvements, whereas in the control group, only receptive skills showed a significant improvement, and the magnitude of improvement in the experimental group was greater than that in the control group. These results demonstrate that the AI-assisted Chinese construction grammar teaching model can effectively enhance learners’ receptive and productive abilities in Chinese construction grammar, whereas the traditional teaching model failed to significantly improve productive abilities.

The findings of this study provide a practical, technology-driven approach and empirical evidence for the teaching of Chinese construction grammar, offering relevant guidance for the design and application of intelligent language teaching systems.

References

1. Wu, Y. (2017). Teaching Chinese as an international language: A Singapore perspective. Cambridge University Press.
2. Sharma, B. K. (2018). Chinese as a global language: Negotiating ideologies and identities. Global Chinese, 4(1), 1-10.
3. Diessel, H., Dabrowska, E., & Divjak, D. (2019). Usage-based construction grammar. Cognitive linguistics, 2(1), 50-80.
4. Zima, E., & Bergs, A. (2017). Multimodality and construction grammar. Linguistics Vanguard, 3(s1), 20161006.
5. Wasserscheidt, P. (2019). Construction grammar: Basic principles and concepts. Ukrainian Linguistics Journal, 49(1), 94-116.
6. Hoffman, T., & Bergs, A. (2018). A construction grammar approach to genre. CogniTextes. Revue de l’Association française de linguistique cognitive, 18(Volume 18).
7. Kim, H., Kim, T. E., & Park, J. (2024). The cognitive construction-grammar approach to teaching the Chinese Ba construction in a foreign language classroom. International Review of Applied Linguistics in Language Teaching, 62(2), 457-483.
8. Lu, X. (2020). Usage-Based Grammar Teaching: The Case of Chinese Verbal Construction" V-lai-V-qu". Journal of Second Language Acquisition and Teaching, 26, 75-98.
9. Jiang, C., & Wen, X. (2022). Constructional network at work in second language acquisition. Asian-Pacific Journal of Second and Foreign Language Education, 7(1), 12.
10. Jing-Schmidt, Z. (2019). Grammatical constructions and Chinese discourse. In The Routledge handbook of Chinese discourse analysis (pp. 102-115). Routledge.
11. Garzón, J., Kinshuk, Burgos, D., & Tlili, A. (2025). Advantages and challenges associated with mobile learning in education: A systematic literature review. Journal of Computers in Education, 12(4), 1173-1205.

12. Li, B. (2022). The application trend and advantages of mobile technology in higher education. Mobile Information Systems, 2022(1), 3874673.
13. Criollo-C, S., Guerrero-Arias, A., Jaramillo-Alcázar, Á., & Luján-Mora, S. (2021). Mobile learning technologies for education: Benefits and pending issues. Applied Sciences, 11(9), 4111.
14. Furió, D., Juan, M. C., Seguí, I., & Vivó, R. (2015). Mobile learning vs. traditional classroom lessons: a comparative study. Journal of Computer Assisted Learning, 31(3), 189-201.

About the Author

Jiayi Yang holds a bachelor's degree in International Chinese Language Education from North China University of Water Resources and Electric Power. Her research interests cover second language acquisition, Chinese rhetorical strategies, stylistic linguistics and international Chinese language education. She has rich frontline teaching experience in Chinese language instruction, and has conducted empirical research on mobile-assisted Chinese teaching models for second language learners, with relevant empirical achievements in digital Chinese education and cross-cultural Chinese communication.

Xu Wang serves as Faculty Supervisor at the School of Data Science, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), with core research interests covering artificial intelligence, big data analytics, generative large models and medical-industrial integration; concurrent appointments include Adjunct Doctoral Supervisor, Committee Expert at the China Academy of Information and Communications Technology, specialist of multiple national research working groups focusing on artificial intelligence and enterprise digital transformation, and committee member of intelligent healthcare, large model and generative AI professional committees under multiple national academic societies, alongside peer reviewer for academic journals and research funding programs sponsored by national and municipal science and technology authorities, competition judge for numerous national innovation and talent selection initiatives, and co-author of authoritative industry reports including the China Data Analytics Industry Annual Development Report (2024), having received multiple national and municipal science and technology achievement awards and honors from national artificial intelligence competitions, with research outputs selected as outstanding exemplary cases of artificial intelligence large models and intelligent technology innovation at authoritative national industry conferences.