Research Proposal

LLM-Driven Multimedia System for International Chinese Education

Contents

1 Introduction

1.1 Background

International Chinese, as a distinctive linguistic paradigm, endeavors to bridge cross-cultural communication gaps and surmount language and cultural barriers through Chinese education. To accomplish this objective, the International Chinese multimedia learning system will synthesize Chinese linguistic knowledge with traditional Chinese culture, thereby bolstering learners' engagement and active involvement. Additionally, the system will incorporate various forms of expression, such as graphics and videos, utilizing multimodal pedagogical strategies. This comprehensive learning system aims to provide robust support for global Chinese learners and assume an instrumental role in propelling research in International Chinese education.

The system capitalizes on a Large Language Model (LLM) to undertake comprehensive training on a voluminous Chinese corpus, with the aim of developing an omnipotent AI assistant for international Chinese instruction, thereby facilitating 24/7 high-caliber customized pedagogy within an incorporating multimedia architecture for learners. Simultaneously, through the integration of multiple speech models, the system will accurately capture learners' spoken pronunciation, swiftly analyze semantics, and generate appropriate responses, facilitating oral interactions across diverse scenarios. The implemented algorithm will compare Chinese learners' pronunciation to the standard pronunciation, ensuring accurate rectification. Moreover, the system will utilize learners' dialogue data for self-learning analysis, intelligently recommend tailored learning content, and match suitable topics and traditional Chinese culture passages, ultimately enhancing the efficacy of Chinese language learning.

The Chinese Proficiency Test (HSK) is an international standardized exam designed to assess the communicative abilities of non-native Chinese speakers in daily life, academic, and professional contexts, covering HSK levels 1 to 6. Our research research will rely on HSK vocabulary and traditional Chinese culture classics to construct a rich Chinese corpus for the International Chinese learning multimedia system. For each topic scenario, the system will provide corresponding dialogue tasks for learners. During the interaction process, the system will autonomously correct learners' grammatical errors in their dialogues. Upon dialogue completion, the system will evaluate learners' responses based on factors such as topic relevance and semantic coherence, providing targeted recommendations for improvement.

Media convergence represents a media development model that amalgamates traditional media, emerging media, and internet technologies. In this research research, we will fully leverage the unique advantages of various media types, presenting International Chinese knowledge and traditional Chinese culture in a multimodal format, making the Chinese learning process more intuitive and comprehensible.

Upon completion, this research research will offer a comprehensive solution ranging from corpus construction and model training to innovative applications. It aims to inherit and promote traditional Chinese culture, embody the spirit of the times, and contribute to the development of an International Chinese learning system characterized by Chinese attributes in the new era of socialism.Upon completion, this research research will encompass a comprehensive solution, spanning corpus construction, model training, and innovative applications, with the objective of preserving and promoting traditional Chinese culture, epitomizing the spirit of the times, and contributing to the development of an International Chinese learning system characterized by Chinese attributes in the new era of socialism.

1.2 Research Objectives

As our country's international standing continues to rise, cultural exchanges between China and foreign nations have become increasingly frequent. The exceptional traditional culture of the Chinese nation has progressively earned international recognition and admiration. Under the scientific perspective of "Chinese fever," the dissemination of international Chinese learning has been unprecedentedly strengthened

In recent years, the core issues concerning teachers, teaching materials, and teaching methods—the "three teachings"—persist as fundamental challenges in international Chinese education and global promotion of the Chinese language. Furthermore, following the COVID-19 pandemic, international Chinese teaching has undergone a transformative shift from offline to online, giving rise to novel demands in the "three teachings" realms:

- For teachers, mastering Chinese language teaching skills is essential, as is familiarity with the operational principles of the internet for effective online instruction.

- Concerning teaching materials, online resources should extend beyond written formats, incorporating graphics, videos, and other multimedia elements to develop multimodal Chinese educational resources.

- With respect to teaching methods, adapting to online teaching characteristics, devising innovative teaching concepts and systems, and prioritizing real-time interaction without overlooking the persistence of cloud-based content are crucial.

To address the "three teachings" challenges, this research will develop a Chinese learning AI assistant using large language models to tackle the "teacher" issue; implement multimodal teaching with multimedia technology to address the "teaching materials" problem; and ultimately, by integrating the aforementioned technologies and combining the multimodal corpus of traditional Chinese cultural classics, construct an LLM-based international Chinese learning multimedia system to resolve the "teaching methods" dilemma.

2 Previous Research

2.1 Status of Domestic Research

On March 30, 2023, the virtual reality (VR) version of the course series "Chinese for Beginners" led by Associate Professor Liu Xiaoyu from Peking University was launched globally on the Coursera platform. Learners can access the course through VR headsets or desktops, immersing themselves in authentic Chinese living scenarios. While receiving real-time feedback, they can practice listening and speaking skills, as well as gain insights into Chinese society and culture.

This course from Peking University successfully integrates international Chinese education with virtual reality (VR), creating multimodal Chinese teaching resources.

However, it lacks real-time online interaction and cannot provide personalized learning strategies and content recommendations based on the characteristics and needs of individual learners. In comparison to the intelligent Chinese teaching AI assistant proposed in this research, it lacks flexibility.
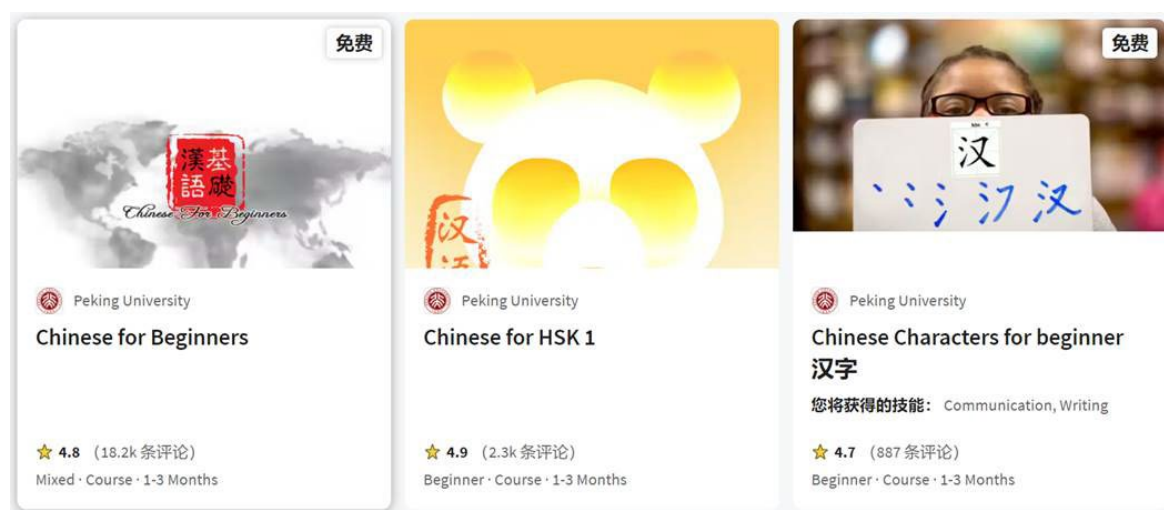

Figure 1 Peking University's International Chinese Course Series

2.2 Status of International Research

Researchers at Michigan State University in the United States have developed a system for automatically assessing students' Chinese writing skills. This system can analyze students' compositions and provide personalized feedback and suggestions.

However, compared to the GPT technology utilized in this research, the intelligence of this system has certain limitations. Furthermore, the scope of its application is restricted to the evaluation of writing skills, lacking versatility. In contrast, the LLM-based international Chinese learning multimedia system proposed in this research offers greater multifunctionality and multimodal teaching capabilities, providing students with more comprehensive support.

2.3 Current Trends in Development

In recent years, the number of Chinese language learners has been steadily increasing, with Chinese education being carried out in more than 180 countries and regions worldwide. Chinese has been incorporated into the national education systems of 81 countries, with over 30 million people learning the language abroad. Chinese teaching is expanding in various aspects, levels, and fields, with diverse and specialized talent cultivation. International Chinese education is shifting from a single focus on Chinese language instruction to more diversified teaching methods and high-quality, content-rich development.

On March 22, 2023, the global education think tank HolonIQ released a report on the Chinese learning market, predicting that by 2027, the global Chinese learning market will double in size. With a compound annual growth rate of 12.1%, it is expected to grow from $7.4 billion to $13.1 billion. Chinese is currently the fastest-growing language learning market globally, with the demand for Chinese learning rapidly increasing. This is especially true for K-12 students within the Chinese diaspora, which consists of approximately 60 million individuals in total.
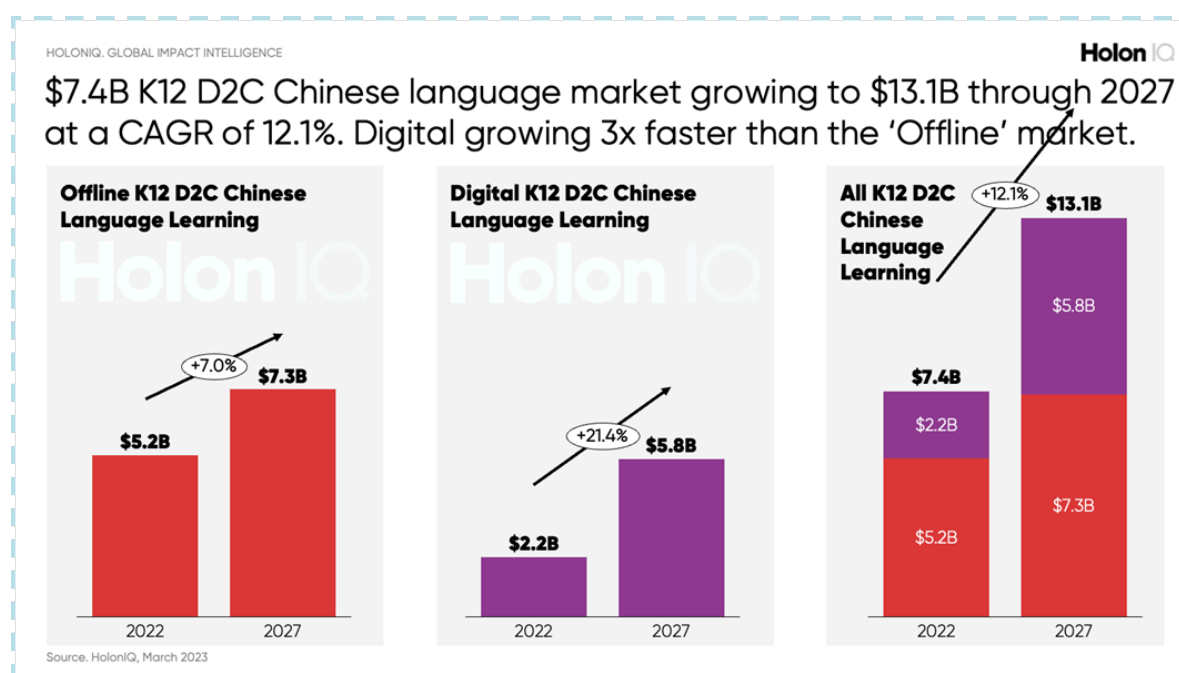
Figure 2 HolonIQ's Report on the Chinese Learning Market

Since the outbreak of the COVID-19 pandemic, the integration of artificial intelligence (AI) technology and the field of education has deepened, especially in foreign language teaching and English learning. Deep learning algorithms have made personalized learning and recommendations possible, and multimedia technology has enabled the creation of multimodal educational materials, attracting widespread attention and application.

The field of international Chinese education has also begun to explore the combination of AI technology and multimedia technology. However, the application of large language models in international Chinese education remains relatively limited, offering significant room for further development.

3 Research methods

3.1 Overview of the Research

The main objective of the international Chinese learning multimedia system is to develop an intelligent multimedia platform for international Chinese learners, utilizing the HSK framework as a benchmark, while also considering cross-cultural communication competencies.
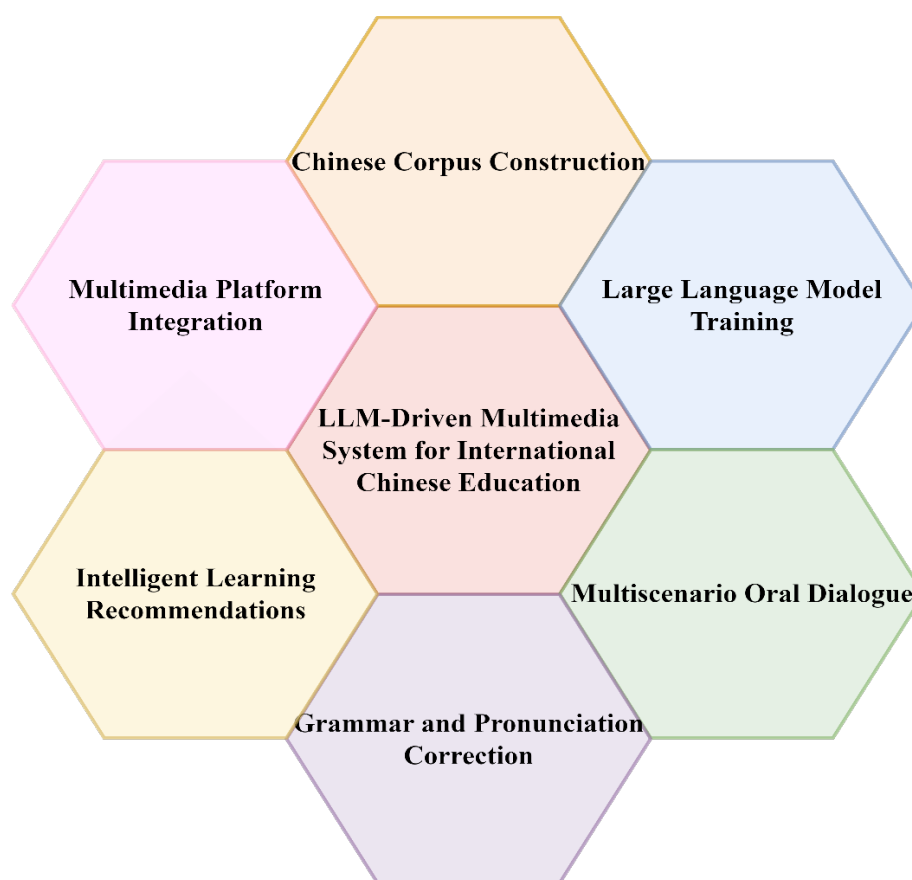

Figure 3 Functional Module Diagram of the System

As a result of the COVID-19 pandemic, international Chinese learning has gradually shifted from offline physical classrooms to online virtual classrooms. Due to the unique temporal and spatial constraints, teacher-student interactions are primarily limited to class time, leading to a lack of engagement and interaction outside of lessons. Consequently, learners often encounter questions without readily available answers. Furthermore, challenges in online teaching include weak cultural exchanges, insufficient oral training, monotonous and uninteresting teaching materials, and a lack of diversity in classroom formats.

3.2 Chinese Corpus Construction

To address weak cultural exchanges, this research aims to construct a multimodal corpus using a diverse range of traditional Chinese cultural texts. The corpus's technical framework comprises a foundational layer, an intermediate layer, and an interactive layer.

The *foundational layer* focuses on collecting and organizing corpus elements, including acquiring, organizing, and archiving multimodal data. This involves gathering images from texts, compiling textual stories from audio and video sources, and designing multimodal data element alignment schemes.

The *intermediate layer* involves constructing algorithmic models and designing the business logic layer. This includes multimodal corpus association analysis and presentation algorithms, search algorithms, and algorithms for sentiment analysis, cross-modal translation, special element location retrieval, and series story association analysis in the context of multimodal corpus integration. Additionally, it encompasses the design of the logic layer for algorithm integration and association.

The *interactive layer* serves as the learner interface, enabling retrieval and dynamic interaction to foster an integrated virtual community for inspection, research, application, and learning. The multimodal corpus employs data processing, storage management, and visualization interaction analysis technologies to facilitate modality-related queries, visual association analysis, and task information discovery under multimodal fusion.

Based on the multimodal corpus of traditional Chinese culture, a substantial amount of Chinese education-related textual materials will be collected and organized, including textbook content, teacher lectures, and exercise explanations. The texts will undergo data cleaning, such as removing special symbols, word segmentation, and part-of-speech tagging, for subsequent model training.

3.3 Large Language Model Training

To address the issue of limited interaction outside of class, this research plans to adopt the large-scale Transformer-based language model GPT-3.5 and employ the PLMF (Pretrained Language Model Fine-tuning) method from transfer learning. The Chinese corpus will serve as the data foundation for the model, with the aim of designing and implementing a web-based Chinese teaching AI assistant. By efficiently generating Chinese vocabulary explanations and learning prompts, the AI assistant is expected to provide 24-hour uninterrupted high-quality Q&A services in the future.

3.4 Multiscenario Oral Dialogue

To address insufficient oral training, we plan to develop AI dialogue assistants for various scenarios using large language model training. These assistants will engage in text-based conversations within specific contexts and simulate oral dialogues on particular topics. We will

train specialized exercise assistants for multiple scenarios using self-learning technologies and a multimodal corpus.

3.4.1 Implementation Details of Self-learning Technology

We plan to integrate natural language processing and deep learning algorithms to achieve self-learning functionality. Our model includes the following key steps:

- *Rating Model Training*: Utilize manually annotated oral training data, covering various difficulty levels and topics. Improve prediction accuracy through supervised learning and transfer learning techniques.

- *Chinese Proficiency Assessment*: Analyze learners' inputs (vocabulary, grammar, and semantic understanding) and assess oral proficiency by comparing inputs with standard answers at various difficulty levels.

3.4.2 Scenario-specific Dialogue Model Training

We plan to train models for different life scenarios, including but not limited to:

- *Travel and Transportation*: Flight booking, hotel check-in, sightseeing tours, etc.

- *Study and Work*: Classroom discussions, job interviews, research presentations, etc.

For each topic scenario, the AI provides dialogue tasks, scores learner performance based on factors such as topic relevance and semantic coherence, and offers targeted improvement suggestions.

3.4.3 Difficulty-based Dialogue Model Training

We divide oral training into three difficulty levels: basic topics, advanced topics, and expert topics. The model uses self-learning technology to assess learners' oral proficiency and match appropriate exercises. Grading criteria reference the HSK syllabus and past students' scores. Specific contents and details are as follows:

- *Basic Topics*: For HSK 1-3 learners, covering daily life scenarios. Training focuses on basic vocabulary, grammar, and pronunciation.

- *Advanced Topics*: For HSK 4-5 learners, covering more complex scenarios. Training emphasizes vocabulary expansion, language expression accuracy, and fluency.

- *Expert Topics*: For HSK 6 and above learners, involving abstract themes and social issues. Training focuses on in-depth thinking, logical reasoning, argumentation skills, and authentic expression.

3.5 Grammar and Pronunciation Correction

3.5.1 Grammar Correction

Grammar correction aims to analyze and rectify Chinese learners' text input. We plan to adopt the following technologies and methods:

- *Dependency Parsing*: Identify grammatical components (e.g., subject, predicate, object) and assess sentence conformity with grammatical rules.

- *Rule-based Correction*: Design a rule set based on Chinese grammar rules, allowing the system to detect and correct learners' grammatical errors.

- *Statistical Language Model*: Train a model using a Chinese corpus to evaluate linguistic regularity; correct sentences based on probability values. Utilize deep learning methods (e.g., Transformer, BERT) combined with annotated corpora to train high-performance grammar correction models.

- *Multi-model Fusion*: Enhance accuracy and robustness by fusing various grammar correction models to develop a comprehensive and superior model.

3.5.2 Pronunciation Correction

For pronunciation correction, we plan to adopt the following technologies:

- *Speech Recognition*: Use Automatic Speech Recognition (ASR) to convert learners' speech input into text for data processing.

- *Prosody Analysis*: Evaluate pronunciation accuracy and naturalness by analyzing prosodic features (e.g., pitch, intensity, speech rate).

- *Rule-based Pronunciation Correction*: Design a set of pronunciation correction rules based on Chinese pronunciation rules and phonological knowledge, enabling the system to detect and correct learners' pronunciation errors.

3.6 Intelligent Learning Recommendations

Addressing the issue of dull and monotonous teaching materials, our future research will rely on HSK vocabulary and traditional Chinese cultural classics to construct a rich Chinese corpus. The system will employ self-learning analysis of learners' dialogue data to intelligently recommend targeted learning content, matching appropriate topics and Chinese traditional cultural passages. This approach aims to enhance the effectiveness of Chinese language learning for learners while maintaining scientific, innovative, rigorous, and academic standards.

3.7 Multimedia Platform Integration

Addressing the issue of monotonous classroom formats, our future research will rely on short video and live streaming platforms such as YouTube and TikTok to construct a multi-terminal online teaching tool for overseas Chinese learners—HSKhelper. This platform will integrate various teaching resources in multimodal formats, such as video, audio, and graphics, providing international Chinese learners with a rich and engaging learning experience. By regularly hosting activities like HSK fun quizzes, oral Chinese demonstrations, and Chinese character composition showcases, we aim to enhance learners' motivation to learn while adhering to scientific, innovative, rigorous, and academic practices.

4 Schedule

4.1 System Feasibility Analysis and Technical Preparation

June 2023 - July 2023, we will conduct a comprehensive analysis to further evaluate the feasibility of the "International Chinese Learning Multimedia System". We will master the necessary technologies and skills for implementing this system, including front-end and back-end web development, Python data processing and mining techniques, TensorFlow deep learning framework, and large language model training techniques.

4.2 System Refinement and Multimodal Corpus Construction

August 2023 - October 2023, we will further refine the functional requirements of the system. To construct a traditional Chinese culture multimodal corpus, we will obtain the required data, initially saved in formats such as Word and Excel. We will design an automated data cleaning algorithm for these documents and develop web crawlers for automatically retrieving content from traditional Chinese cultural classics.

4.3 Design and Implementation of Chinese Learning System

November 2023 - February 2024, we will implement the "International Chinese Learning Multimedia System", including training the large language model, setting up the software development environment, designing and implementing the deep learning-based grammar and pronunciation models, and intelligent recommendation algorithm, as well as system software architecture design, software coding and testing, and phase evaluation.

4.4 System Integration, Comprehensive Optimization, and Final Testing

March 2024 - April 2024, we will rigorously examine the preliminary "International Chinese Learning Multimedia System" by focusing on final integration of sub-modules, such as the AI language assistant, speech recognition models, and content recommendation algorithms. Additionally, we will optimize the system's performance, addressing bottlenecks and enhancing the user interface, while conducting extensive tests covering functionality, reliability, and compatibility.

4.5 Finalizing Documentation and Preparing for Research Defense

May 2024 - June 2024, we will meticulously compile and prepare comprehensive documentation, including detailed descriptions of the system's design, implementation, and testing procedures. We will refine the software, incorporating feedback from tests and user experiences, to ensure a polished release. Concurrently, we will prepare a robust research defense, highlighting the scientific, innovative, rigorous, and academic aspects of our work.

5 Expected Results and Conclusions

Upon completion of our project, we expect to make significant advancements in the field of Chinese language learning by addressing critical challenges and delivering valuable results:

- **Enhanced After-Class Interaction**: We will develop a web-based Chinese teaching AI assistant to improve after-class interaction, providing round-the-clock, high-quality Q&A services through efficient vocabulary definitions and learning prompt generation.

- **Innovative Oral Training Solutions**: We will integrate various speech recognition models to tackle inadequate oral training, enabling real-time dialogue semantics analysis and high-quality response generation. This approach will facilitate AI-assisted correction and instruction of spoken language.

- **Engaging Multimedia Content**: We will employ multimedia platforms, such as YouTube, to overcome the monotony of traditional classroom formats. By disseminating graphic and video tutorials that incorporate animations and traditional cultural stories, we will create a more engaging learning experience.

- **Intelligent Content Recommendation**: We will construct a corpus combining HSK vocabulary and Chinese cultural classics to address unengaging teaching materials. By annotating corpus data based on difficulty and context, we will enable intelligent recommendations of multimodal content, enriching the learning process.

Upon project completion, we expect the following outcomes:

- A comprehensive International Chinese Learning Multimedia System for effective language learning.

- High-precision intelligent recommendation and spoken dialogue models for personalized experiences.

- Cultivation of team members' innovation, cross-cultural communication, and collaboration skills.

- Securing software copyrights to protect associated intellectual property.

- Publication of 1 to 2 academic papers, showcasing our findings and contributions.

These anticipated results will lay a solid foundation for future advancements in the realms of language learning and multimedia technology, catalyzing progress and innovation in the field. The system offers an all-encompassing solution, encompassing corpus construction, model training, parameter optimization, innovative applications, and promotional efforts. The primary objective is to preserve and propagate the rich traditional Chinese culture while embracing the spirit of contemporary times, thereby contributing to the development of an international Chinese learning system that embodies the unique characteristics of socialism in the new era. This holistic approach will further reinforce the platform's pivotal role in pushing the boundaries of language learning and multimedia technologies, ultimately benefiting learners across the globe.

6 References

[1] 王晓彤(主编). (2019).大学英语学术写作.北京:中国人民大学出版社.

[2] 刘妍.(2023).人工智能技术对国际中文教育的影响与对策——在 ChatGPT 出现的背景下. 中国现代教育装备(09),16-18. doi:10.13492/j.cnki.cmee.2023.09.035.

[3] 胡云晚.(2023).国际中文教育之科技汉语研究四十年.贵州师范大学学报(社会科学版)(03),60-71. doi:10.16614/j.gznuj.skb.2023.03.006.

[4] 刘利,周小兵,高雪松,潘海峰,刘晓海,饶高琦... & 韩晓明.(2023)."ChatGPT 来了：国际中文教育的新机遇与新挑战"大家谈(上). 语言教学与研究(03),1-14.

[5] 王乐韬&蒋跃.(2023).在线汉语国际教育的发展策略研究. 教育理论与实践(12),20-23.

[6] 包昊罡 & 王轶晰.(2023).ChatGPT 背景下的基础教育数字资源建设. 中国基础教育(05),39-42.

[7] 陈虹羽 & 李东伟.(2023).基于案例分析的国际中文教育线上教学问题及解决对策研究. 华北理工大学学报(社会科学版)(03),74-78+85.

[8] Firat, M. (2023). How chat GPT can transform autodidactic experiences and open education. Department of Distance Education, Open Education Faculty, Anadolu Unive.

[9] Milano, S., McGrane, J. A., & Leonelli, S. (2023). Large language models challenge the future of higher education. Nature Machine Intelligence, 1-2.

[10] Jeon, J., Lee, S., & Choi, S. (2023). A systematic review of research on speech-recognition chatbots for language learning: Implications for future directions in the era of large language models. Interactive Learning Environments, 1-19.

[11] Kwon, T. (2023). Interfaces for Personalized Language Learning with Generative Language Models (Doctoral dissertation, Columbia University).

[12] Laato, S., Morschheuser, B., Hamari, J., & Björne, J. AI-assisted Learning with ChatGPT and Large Language Models: Implications for Higher Education.

[13] Peng, B., Li, C., He, P., Galley, M., & Gao, J. (2023). Instruction tuning with gpt-4. arXiv preprint arXiv:2304.03277.

[14] Gong, T., Lyu, C., Zhang, S., Wang, Y., Zheng, M., Zhao, Q., ... & Chen, K. (2023). Multimodal-gpt: A vision and language model for dialogue with humans. arXiv preprint arXiv:2305.04790.

[15] Anand, Y., Nussbaum, Z., Duderstadt, B., Schmidt, B., & Mulyar, A. (2023). Gpt4all: Training an assistant-style chatbot with large scale data distillation from gpt-3.5-turbo. GitHub.