



**LANYU: LEVERAGE GENERATIVE ARTIFICIAL INTELLIGENCE IN
ASSISTING FOREIGN LANGUAGE LEARNING**

Lappeenranta–Lahti University of Technology LUT

Bachelor's Programme in Software and System Engineering, Bachelor's thesis

2024

Chuanzhang Chen

Examiners: Assistant Professor Dominik Siemon

Dr. Shouhua Zhang, Ph.D

ABSTRACT

Lappeenranta–Lahti University of Technology LUT

LUT School of Engineering Sciences

Software and Systems Engineering

In co-operation with partner university: Hebei University of Technology (HEBUT)

Chuanzhang Chen

LanYu: Leverage Generative Artificial Intelligence in Assisting Foreign Language Learning

Bachelor's thesis

2024

42 pages, 12 figures, 4 tables and 1 appendix

Examiners: Assistant Professor Dominik Siemon and Dr. Shouhua Zhang, Ph.D

Keywords: Design science research, generative artificial intelligence, large language model, foreign language learning

Nowadays, mastering one or more foreign languages has become the norm. Especially with the expanding economic and cultural influence of China, there is a growing interest in learning Chinese. However, due to the limitations of traditional foreign language teaching methods and the shortage of appropriate and up-to-date learning materials, learners fail to gain ideal learning outcomes and gradually decrease learning enthusiasm. At the same time, this situation also places a burden on educators. The advent of generative artificial intelligence (GAI) has opened up new avenues for foreign language learning, yet there is a lack of clear design principles to guide its application in this field. This study adopts a design science research (DSR) methodology to derive three practical design principles for applying GAI in assisting foreign language learning, leading to the development of the *LanYu* system. This system leverages GAI to allow users only to enter Chinese characters or phrases instead of prompts, and obtain two contextualized dialogues reflecting real-life scenarios, along with corresponding vivid visual illustrations. Feedback from user studies demonstrates that this system efficiently assists learners in learning Chinese by not only providing professional and easy-to-understand learning materials but also offering an enjoyable learning experience. This research provides design guidance for incorporating GAI into foreign language learning and lays the foundation for establishing an innovative paradigm in foreign language education in the future.

ABBREVIATIONS

GAI	Generative artificial intelligence
AI	Artificial intelligence
LLM	Large language model
T2I	Text to image
DSR	Design science research
DR	Design requirement
CoT	Chain of thought
TAM	Technology acceptance model

Table of contents

Abstract

Abbreviations

1	Introduction.....	6
2	Related Work.....	9
2.1	GAI for Foreign Language Learning.....	9
2.2	Prompt Engineering.....	10
2.3	Consistency in Image Generation.....	10
3	Methodology.....	12
3.1	Research Approach.....	12
3.2	Identification of Problems and Motivations.....	14
3.3	Objectives of a Solution.....	15
3.4	Design and Development.....	17
3.4.1	Derivation Design Principles.....	17
3.4.2	LanYu System.....	19
4	Demonstration and Evaluation.....	26
5	Discussion and Limitation.....	30
5.1	Discussion.....	30
5.1.1	Design Principles.....	30
5.1.2	LanYu System.....	31
5.1.3	Consistency in Image Generation.....	32
5.2	Limitation.....	33
5.2.1	User Interface Design.....	33
5.2.2	Problems with Unknown Characters in Generated Images.....	33
5.2.3	Delivering Multimodal Output.....	34
6	Conclusions.....	36
	References.....	37

Appendices

Appendix 1 Questionnaire

1 Introduction

“One who knows nothing about a foreign language knows nothing about oneself,” said Goethe, clarifying the importance of learning a foreign language (Ratcliffe, 2018). In an era of increasing international communication, learning one or more foreign languages strengthens the development of various personal skills and offers potential advantages for career prospects and academic achievement (Fox, Corretjer and Webb, 2019). Hence, foreign language education has been widely prioritized. According to the statistics on foreign language learning reported by Eurostat, in 2021, more than half of EU students (60.6%) in lower secondary education were learning two or more foreign languages, with Finland leading at 98.4%, the highest percentage among all EU Member States (Eurostat, 2023). In upper secondary education, approximately half of EU students (49.5%) studied at least two kinds of foreign languages, particularly in Romania, where almost the entire student population in upper secondary education (98.0%) studied multiple foreign languages, and in Finland (87.5%) and Luxembourg (82.9%) where the proportion exceeds 80% (Eurostat, 2023). This trend also extends to higher education institutions, as numerous universities offer language courses covering various foreign languages, particularly Chinese courses, which have gained popularity due to the rapid development of China (H. Zheng and L. Zheng, 2020; G. X. Zhang and L. M. Li, 2010).

The traditional pattern of foreign language education mainly centers around the teacher, with the classroom serving as the primary learning environment, demands students attend classes regularly and mostly learn in the classroom (Collins and Muñoz, 2016). It usually lasts several weeks or months and requires strict adherence to pre-determined lesson plans and curricula. However, this approach tends to be overly rigid, overemphasizing the memorization of grammatical structures and vocabulary (Renau and María, 2016). In the long run, such a stereotyped teaching model will diminish students’ motivation and interest in foreign language acquisition and limit their creative ability. Preparing learning materials is another challenge in traditional foreign language teaching. On the one hand, teachers must exhaust a great deal of time and effort into searching and organizing appropriate textual and visual learning resources to enrich learning content, placing a considerable burden on them. On the other hand, students’ textbooks and learning resources are often outdated and overly standardized in content, deviating from real-life contexts, making it difficult for students to apply acquired knowledge in their daily interactions and potentially discouraging their motivation to learn (G. X. Zhang and L. M. Li, 2010).

Nowadays, incorporating advanced technology to assist foreign language teaching has become a modern trend. Golonka et al. (2014) conducted a comprehensive review of early

technological applications in foreign language education, illustrating the positive impact that technology can have on language learning and teaching. Building upon this, Shadiev and Yang (2020) compared the applications of early and modern technologies in foreign language education. Covering relevant literature between 2014 and 2019, they provided a more extensive analysis of twenty-three distinct technologies used in foreign language learning, further proving that emerging technologies not only effectively provide foreign language learning support but also increase the motivation and interest of learners.

With the advent of ChatGPT in 2022, generative artificial intelligence (GAI) has sparked a revolution that profoundly impacts every aspect of life (OpenAI, 2022). This is a new type of artificial intelligence (AI) that not only has the ability to understand user input and respond, but can also adapt to personal preferences to provide personalized answers in diverse forms, such as text, images, video, etc. (Touvron et al., 2023; Nichol et al., 2021; Saharia et al., 2022; Ho et al., 2022). Among the most representative of these, ChatGPT, a conversational large language model (LLM), has abilities to produce human-like text based on context, which has a broad application foreground in healthcare, business and finance, law, education, and scientific research (Ray, 2023). Through a series of iterative updates and improvements by OpenAI, ChatGPT has evolved from the GPT-2 and GPT-3 models and is now the latest GPT-4 model (Radford et al., 2019; Brown et al., 2020). This advanced model outperforms most previous LLMs in understanding a wide range of languages and exhibiting human-like capabilities in a variety of academic and professional exams (Achiam et al., 2023). Another notable advancement is the DALLE-E 3, a text-to-image (T2I) generation model developed by OpenAI, which allows users to produce vivid images solely from textual descriptions, outperforming other models in various assessments (Betker et al., 2023).

GAI has demonstrated its exciting potential, and its practical applications are being widely explored. However, despite these advances, the practical application of GAI to specific scenarios, particularly in foreign language learning, still lacks clear design principles. Given the current situation, this paper proposes a research question (RQ) to investigate the potential design principles for more effective incorporation of GAI to assist foreign language learning:

RQ: What design principle should be followed when leveraging GAI in assisting foreign language learning?

Therefore, this thesis aims to investigate this RQ and explore appropriate design principles by adapting the design science research (DSR) approach (Hevner et al., 2004; Peffers et al., 2007; Möller, Guggenberger and Otto, 2020). The chapters are organized as follows: Chapter 2 reviews the application of GAI in foreign language learning and discusses related tech-

niques for optimizing GAI-generated content. In Chapter 3, the paper outlines the process of deriving the design principles based on the DSR methodology. Through user interviews and a literature review, this study has successfully identified essential design requirements (DRs) and derived three innovative design principles. These principles are further used to guide the development of LanYu (澜语). This prototype system utilizes samples collected from the textbook to fine-tune the LLM and couple it with the T2I generation model, allowing the system to generate professional and closely aligned real-world contextual conversations based on words or phrases the user enters. In addition, by integrating textual and visual illustrations, the system provides users with an immersive learning experience that promotes user engagement and effectively assists them in grasping and gaining a deeper understanding and impressions of vocabulary in real-world contexts. Chapter 4 presents the results and user feedback from a user study. The results validate that adhering to the developed design principles in guiding the development of GAI for foreign language learning applications not only provides professional and attractive learning resources to bolster their learning but also enriches the user experience. Finally, the paper summarizes the research's primary contributions, highlights existing limitations of this research, and suggests opportunities for improvement as well as directions for future exploration.

2 Related Work

2.1 GAI for Foreign Language Learning

GAI has demonstrated its potential to support students in learning foreign languages through its advanced semantic understanding and capacity to interact with users and continuously provide personalized and real-time responses. With the support of GAI, students can gain valuable learning resources, improve their writing skills, and increase their motivation to practice when learning a foreign language, such as English (Young and Shishido, 2023; Kim, Shim and Shim, 2023; C. Song and Y. Song, 2023). Moreover, the application of GAI to foreign language learning is not limited to English; it has also been shown to be effective in supporting the learning of other foreign languages, such as German, Spanish, Maltese, and so on (Athanassopoulos et al., 2023; Tarp and Nomdedeu-Rull, 2024; Żammit, 2023).

Additionally, there is a growing interest in learning Chinese due to China's rapid development. According to a report by the Chinese Ministry of Education, over 30 million individuals studied Chinese outside of China until 2023, bringing the total number of Chinese learners and users close to nearly 200 million (Ministry of Education of the People's Republic of China, 2023). However, learning Chinese presents plenty of challenges, such as pronunciation, grammar, Chinese characters, and various other difficulties. As a result, there has been an increased interest in applying GAI to support Chinese language learning. Recently, a study introduced ChatGPT into the classroom as an instructional tool for Chinese language teaching and learning, providing personalized support for students with varying language proficiency based on different curriculum objectives (J. Li et al., 2023). This study highlighted the ability of ChatGPT to effectively cater to the needs of students at levels of language proficiency and to support a range of teaching and learning activities. While an increasing number of studies are exploring GAI as an effective tool to assist foreign language learning, it is essential to note that GAI is still in the initial stage of development.

Therefore, clear and specific design principles are urgently needed to guide developers and designers to effectively incorporate GAI to develop applications to assist foreign language learning. At present, there are only a limited number of general design principles for GAI applications (Weisz et al., 2024). However, these principles have not yet been refined and explicitly targeted to address the need to integrate GAI into foreign language learning. Therefore, this study aims to explore the design principles of GAI applicable to foreign language learning, thus expecting to fill the research gap in this area.

2.2 Prompt Engineering

LLMs generate responses based on contextual input, but various factors such as prompt structure, length, and content affect the quality of their output (Lu et al., 2022; Kaddour et al., 2023). Prompt engineering represents an effective approach of carefully designing and refining the prompts that are entered to ensure that LLMs better understand the needs of users and generate a response or content that more closely matches those needs (Chen et al., 2023). Effective strategies in prompt engineering include delivering clear instructions, setting specific roles suitable for the task, using special symbols, and continuously trying and adjusting the prompts (Chen et al., 2023). In practice, using prompt engineering techniques to provide foreign language learning resources could not only ensure the quality of the generated learning content but also reduce the burden on the instructor (Lee et al., 2023).

The chain-of-thought (CoT) approach markedly improves the performance of the LLMs on complex reasoning tasks and also ensures that the reasoning process is clear and understandable to the user by imitating the human reasoning process (Z. Zhang et al., 2023). The CoT approach is mainly divided into zero-shot CoT, which guides the model through step-by-step reasoning using phrases such as “*Let’s think step by step,*”, and few-shot CoT, which better meets the user needs by adjusting the response based on provided specific example. Although LLMs are quite knowledgeable, they still may give unreliable or biased answers due to their limitations in understanding the complex laws of the natural world and their inability to recognize unanswerable questions (Nazir and Z. Wang, 2023; Yin et al., 2023). Introducing external expertise and combining it with the CoT approach will significantly improve the reasoning capabilities of LLMs, thereby increasing the accuracy and reliability of the responses they generate (J. Wang et al., 2023). Therefore, this research fine-tunes the LLM by combining dialogues from a textbook few-shot CoT approach to enable students to learn foreign languages more effectively and engagingly.

2.3 Consistency in Image Generation

By simply describing ideas in natural language and combining T2I generation models such as stable diffusion and DALL-E 3, users can easily produce and obtain vivid, high-quality images that dramatically unleash creativity and imagination (Betker et al., 2023; Rombach et al., 2022). In practice, however, users often expect to receive a series of images at once, rather than just a single image, that are all consistent in style and characteristics to ensure a coherent experience, which remains a significant challenge for T2I generation models (Avrahami et al., 2023).

Current research utilizes a variety of methods to ensure consistency in image generation.

For instance, providing sample images beforehand or fine-tuning the parameters in T2I generation models are proven techniques for guaranteeing the coherence of generated images (Ruiz et al., 2022; Han et al., 2023). In a recent study, consistent characters can be generated from text descriptions without providing any target images (Avrahami et al., 2023). All of these methods provide effective, fine-grained control over the image generation process, ensuring that the generated images are consistent in style and characteristics. However, the drawbacks of these approaches are their heavy reliance on computational resources and potentially a long wait for image generation.

Another approach is to use carefully crafted prompts. By describing in detail the style, color, scene, character style, and emotional tone of the image, as well as information such as seed, consistency in the generated images can also be ensured (JoshGreat, 2023). Additionally, guiding the layout of an image through prompts is a simple but effective strategy for maintaining consistency in image style and content (Heidrich and Schreiber, 2023; Mao, 2023). By first describing the overall image style and character image in the input prompt and then further describing each layout in detail, it can also effectively maintain the same visual style and characteristics for the generated image in each layout, thus providing users with a consistent visual experience. While the text prompts are more convenient for guaranteeing image consistency, it is necessary to note that the user first needs to input appropriate prompts. In other words, without relevant technical background knowledge, the user is required to invest time in learning how to design appropriate prompts in order to ensure image consistency.

Therefore, this study proposes an innovative automated image generation approach that allows users to efficiently obtain consistent images in both style and characters that match textual descriptions without requiring background knowledge. Users can achieve this solely through natural language to create clear, attractive, and easy-to-understand visual content that deepens understanding and memorization of characters and phrases, thus assisting students in learning foreign languages better without requiring massive computational resources.

3 Methodology

This section presents in detail how the DSR methodology is implemented step by step to extract DRs, starting from the literature review and user interviews. Then, the design principles are further derived, and the principles are instantiated in a prototype system to answer the RQ mentioned in the introduction.

3.1 Research Approach

Design principles are abstract forms of prescriptive knowledge that support designers and researchers in making decisions during the design process and provide professional guidance in developing artifacts (Gregor, Chandra Kruse and Seidel, 2020; Chandra Kruse, Puroo and Seidel, 2022). It can be categorized into three main types: design principles related to user activity, design principles related to an artifact itself, and design principles combining the characteristics of user activities and the artifact (Gregor, Chandra Kruse and Seidel, 2020). The first type explains how a user uses the artifact, while the second type defines an artifact's essential characteristics, such as its shape, structure, and functions. The third type combines the first two elements by specifying what users are expected to do with the artifact and its features.

In order to derive the design principles, the DSR approach is widely utilized to formulate appropriate design principles to address the challenges and requirements of users. Hevner et al. (2004) initially proposed the DSR approach contained two main components: design and evaluation. This approach aims to integrate and apply existing knowledge to devise innovative solutions to current problems while facilitating the creation of new design knowledge with potential broader impact.

Furthermore, Peffers et al. (2007) introduced an innovative DSR process consisting of a six-stage methodology to guide design science research. This approach begins with problem identification, followed by establishing research objectives, formulating specific requirements based on these objectives, directing the design and development phases, and demonstrating the problem-solving capabilities of artifacts through experiments, simulations, case studies, etc. In addition, the artifacts are evaluated by comparing their actual performance with the relevant techniques, including functional comparisons, performance metrics, and user feedback, among other forms, to decide whether to improve the effectiveness through iterative processes continuously or to defer improvements to future projects. Finally, researchers communicate the importance of the problem, the utility and innovation of the artifacts, and the significance of the design to the relevant academic and practice areas through scholarly publications and professional communication channels.

Recently, Möller, Guggenberger and Otto (2020) categorized design principles into two groups: one that is supportive in justifying future design decisions and another that is reflective artifacts have been developed, leading to further refinement and elaboration of the DSR methodology. They suggest that researchers use a combination of theory, literature review, interviews, or other relevant data sources to infer meta-requirements. In addition, they advocate using mapping diagrams to intuitively understand the connections and internal logical relationships between design principles and meta-requirements.

Given the exciting opportunities presented by GAI applications in foreign language learning and the lack of clear design principles to guide them, this research adapts the DSR methodology to propose a DSR process consisting of six stages, as illustrated in Fig. 1. Specifically, the following sections identify the current challenges in language learning and the use of GAI as a learning tool (Section 3.2) through the literature review. Based on these challenges, expert and user interviews are conducted (Section 3.3) to acquire more concrete DRs and formulate design principles (Section 3.4). Furthermore, these principles are instantiated and applied to a novel system, LanYu (澜语) in Section 3.4. LanYu is an innovative system that leverages GAI to assist foreign language learning, specifically for Chinese, which can generate dialogues in real-world contexts combined with illustrations for contextual understanding, enabling users to comprehend the context of dialogues in real-world scenarios intuitively. Finally, in Chapter 4, experts and users are re-invited to evaluate the system through the technology acceptance model (TAM) (Davis, 1989; Pal and Vanijja, 2020; Ibrahim et al., 2017).

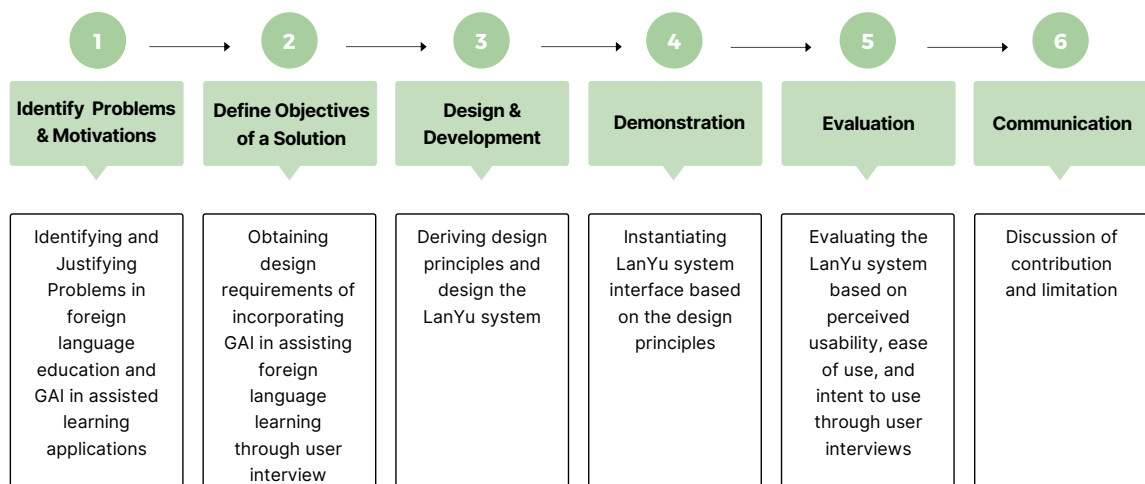


Figure 1. The overview of research approach based on Hevner et al. (2004), Peffers et al. (2007), and Möller, Guggenberger and Otto (2020).

3.2 Identification of Problems and Motivations

Based on the DSR approach, this section discusses the issues and research motivations faced in foreign language education (FE), especially Chinese and GAI in assisted learning applications (GA), as shown in Table 1. The LE highlights the challenges currently present in educational models and the specific difficulties encountered in learning Chinese, while the GA reveals the current limitations of GAI in assisting foreign language learning. Analyzing these two facets can provide an in-depth understanding of the existing issues and further explore practical solutions.

The limitations of the traditional foreign language education pattern have been discussed in the previous introduction and related work sections, emphasizing that current teaching methods are rigid and reduce student motivation (Renau and María, 2016) (**FE1**). These findings demonstrate the urgent need to integrate innovative educational tools and techniques to spark their enthusiasm for learning foreign languages. Additionally, there is a growing interest in learning foreign languages, especially Chinese (Ministry of Education of the People's Republic of China, 2023). However, teachers also face the challenge of accessing learning resources appropriate for students at different learning stages, which requires extensive time and effort in organizing and adjusting these teaching materials to meet the specific needs of students (G. X. Zhang and L. M. Li, 2010; Yue, 2017) (**FE2**). Consequently, it is essential to explore methods or technologies for delivering more abundant and appropriate learning materials to reduce the burden on teachers and enable students to learn more effectively.

The complexity of foreign language learning lies in the diversity of structures and expressions in the language itself, especially in Chinese. The complex grammatical structure of Chinese, its varied pronunciation and tones, and the intricate composition of characters pose multiple challenges for learners (Hu, 2010). Non-native language learners, in particular, always struggle when learning Chinese characters due to the large number of characters and multiple variations in both the writing and pronunciation of each character (M. Li, 2020). Learners are required to spend a huge amount of time and effort to understand the meanings of Chinese characters and phrases and remember the correct writing order and pronunciation rules, resulting in frustration and discouragement. Therefore, finding an efficient method to assist learners in understanding and memorizing these characters and phrases is crucial instead of relying solely on rote repetition (Gong, X. Gao and Lyu, 2020) (**FE3**).

Incorporating GAI into the learning process opens up new opportunities for educators and also brings challenges and issues. Several studies raised concerns about the validity and reliability of GAI-generated responses, which can be incorrect or biased (Javaid et al., 2023; Yu and Guo, 2023) (**GA1**). Furthermore, students may unconsciously absorb and apply such

biased and incorrect information, mistakenly forming incorrect concepts and bringing negative consequences. Thus, ensuring the accuracy and professionalism of generated content is crucial. In addition, differences in the prompts entered into the integrated GAI system may also lead to considerably different quality of the output content, affecting learning outcomes (Chen et al., 2023) (**GA2**). Therefore, applications integrated with GAI are expected to allow users to immediately and effortlessly obtain high-quality responses that meet their needs without constructing complex prompts or possessing technical knowledge.

Table 1. Identified issues from application in language education and GAI in assisted learning applications

No.	Language Education (LE) & GAI in Assisted Learning Applications (GA)
FE1	Current teaching methods are rigid and reduce student motivation (Renau and María, 2016)
FE2	Lacking appropriate learning resources (Yue, 2017)
FE3	The complexity of the Chinese makes it hard to master (Gong, X. Gao and Lyu, 2020)
GA1	Generated responses may be inaccurate or biased (Yu and Guo, 2023; Javaid et al., 2023)
GA2	The quality of prompts directly impacts generated content (Gong, X. Gao and Lyu, 2020)

3.3 Objectives of a Solution

To further obtain user DRs (Table 3) and derive the design principles suitable for applying GAI to foreign language learning, a series of face-to-face interviews were conducted. The participants include two lecturers responsible for teaching Chinese courses (*L1*, *L2*) and three students currently enrolled in these courses (*S1*, *S2*, *S3*), serving as potential users for the interviews. Lecturers are selected based on the following criteria: they are currently engaged in teaching Chinese courses and have rich experience in this field; the students are chosen based on their current enrollment in Chinese courses and previous learning experience.

Each interview lasts about 30-40 minutes, aiming to better understand the requirements of participants and experiences. Before the interviews commence, participants are given a thorough overview of the purpose of this interview and the process to guide them through the session. The interviews primarily revolved around three themes: Previous experience, inquiring about the teaching or learning background of the participants; Current challenges, discussing the difficulties encountered in learning Chinese, such as the difficulty of the Chinese characters, the quality and accessibility of learning resources, and motivation to learn; Opinions on GAI, seeking perspectives on the potential use of GAI to address existing issues in teaching and learning Chinese. The interviewee's information is shown in the Table 2.

Previous learning or teaching experience: The interview began by discussing their previous experiences teaching or learning Chinese, focusing on their motivation and interest. Pre-

Table 2. Interviewee information for obtaining user design principles

ID	Gender	Age	Learning & Professional Experience	Response for & Enrolled Course
L1	Female	38	University lecturer responsible for teaching Chinese for almost 10 years	Basic Chinese & Intermediate Chinese
L2	Female	33	University lecturer with more than one year of experience in teaching Chinese and previously assisted in teaching Chinese	Basic Chinese & Intermediate Chinese
S1	Female	22	Student studied Chinese for at least three years, including elementary school, across eight languages	Basic Chinese & Intermediate Chinese
S2	Male	25	Student studied Chinese for three years, across two languages	Basic Chinese & Intermediate Chinese
S3	Female	20	Student studied Chinese at least three years, including high school, across four languages	Basic Chinese & Intermediate Chinese

vious research highlighted findings from a literature review that suggested a gradual decline in student motivation (**LE1**) while noting an increased interest in learning Chinese, which all interviewees recognized. Furthermore, Lecturers *E1*, *E2*, and Student *S2* suggested that the content in the textbook should be made more appealing, including using rich visual elements such as images and videos, as well as a variety of cases, to increase the motivation of students (**DR1**). Lecturer *E2* also mentioned that it was difficult to find high-quality online resources that were closely aligned with the course content. Lecturer *E1* indicated that searching from the internet for appropriate learning materials was often time-consuming and inefficient, and thus wondered if a more effective way of accessing resources to support Chinese learning could be developed (**DR2**).

“...I feel that the textbooks are not intriguing enough and do not provide sufficient resources for students to learn Chinese.” - E1

“...I am facing challenges in searching for resources on the internet; it is inefficient and time-consuming to get to the one I want” - E1 & E2

Challenges of Chinese learning: In discussing the challenges in learning Chinese, Lecturer *E1* mentioned that students had difficulties in understanding and using phrases in real-life contexts, highlighting that the need to acquire additional resources that combined visual elements, such as comic strips with text, would help students understand and memorize these phrases (**DR3**). Lecturer *E1* also suggested that providing short dialogues based on specific contexts as examples could guide students in using these phrases in real-life situations (**DR4**). Students *S1*, *S2*, and *S3* expressed similar views and a desire to develop a tool that would allow them to type specific phrases or grammar points that would immediately generate images containing detailed explanations and example sentences. Lecturer *E2* also pointed out a significant difficulty in learning Chinese: polysemy. She explained that a single Chi-

nese character or word might convey various meanings depending on its context, and this characteristic brings additional challenges to learning Chinese. For example, the Chinese character “会” can be interpreted as “being able to” in one scenario and “being likely to” in another.

“It is helpful to let students understand the meaning of words and know how to use these phrases by providing examples in both visual and textual in real-world contexts” - E1 & E2

Attitudes and Expectations of GAI: All interviewees expressed a positive willingness to try out GAI. Students *S1*, *S3*, and Lecturer *E1* mentioned that the current GAI might produce some unexpected content and that developing a system or platform specifically for foreign language learning could be beneficial for both teachers and students (**DR5**). However, Lecturer *L1* and *L2* expressed concern about the possible inappropriate expression of GAI-generated content and emphasized the need for ensuring that the content generated by GAI is accurate, professional, and reliable (**DR6**). Similarly, Student *S1* expressed a worry about the possibility of including inappropriate meanings in GAI-generated output.

“...The results it generates are sometimes incorrect or negative, leading to student frustration. If there is a way to make the generated results more professional, then I think it would be more efficient and would not frustrate the student.” - L2

Table 3. Extract design requirements from user interviews and literature review

No.	Design Requirements
DR1	Incorporating images, video, and other multimedia into learning materials to increase the interest and interactivity of a course, thereby motivating students to learn
DR2	Allowing teachers more efficient access to instructional resources also ensures that students have access to high-quality materials that are closely aligned with course objectives and content
DR3	Combining visual elements, such as cartoons and text, to support students’ comprehension and memorization of Chinese characters and phrases
DR4	Provide short, contextualized dialogues as examples to guide students in using Chinese characters and phrases in a real-world context.
DR5	Development of customized GAI tools specific to foreign language learning
DR6	When using GAI for educational purposes, the generated content must be accurate, professional, and reliable in order to avoid conveying inappropriate or incorrect information.

3.4 Design and Development

3.4.1 Derivation Design Principles

In the previous chapters, a literature review has successfully identified the main issues and challenges currently faced in language education using GAI to assist learning and clarified the research motivation. In addition, specific design principles have been collected through

user interviews that demonstrate the basic expectations of actual users when using GAI for foreign language learning. This section derives a set of design principles from these issues and requirements. These principles not only serve as guidelines for developing future GAI applications but also aim to ensure that the applications developed are more effective in assisting foreign language learners in actively achieving their learning goals. The specific design principles are presented below.

User feedback has emphasized the desire for learning materials in diverse formats such as images, text, and video (**DR1**). Moreover, incorporating visual elements with contextualized dialogues to present language use in vivid, real-world scenarios (**DR3, DR4**) can not only facilitate users to comprehend the nuances of phrases in diverse contexts, but also deepen their impression of language use across different situations, significantly increasing their motivation and engagement in learning (**FE1, FE2**). Therefore, this research proposes:

DP1: Principle of Content: When designing and developing foreign language assistance applications incorporating GAI, developers should fully leverage the capabilities of GAI to create diverse, real-world context examples and content. The generated content should be closely aligned with the actual needs of users, helping them understand meanings and enhancing learning motivation.

The literature review (**AF1**) shows that the results produced by the current GAI are not always reliable. User feedback suggests that they are concerned about the professionalism and accuracy of the current GAI results. They also expect the GAI to be trained through highly relevant, professional, and reliable resources to produce professional learning materials to help students cope with language learning difficulties and challenges. (**GA1, FE3, and DR6**). The following design principles are therefore proposed:

DP2: Principle of Trustworthy: When designing and developing foreign language learning applications that incorporate, the GAI should be trained with professional and reliable resources to generate professional, high-quality, and accurate responses, thereby ensuring that the application is effective and trustworthy.

When interacting with GAI, the user is required to have a certain level of technical knowledge to obtain high-quality responses, and users may sometimes need to perform multiple interactions to achieve the desired information (**GA2**). Therefore, it is essential to develop customized GAI tools for foreign language learning that provide high-quality, highly relevant instructional materials based on specific learning goals and course content (**DR5, DR2**). Design principles for collaborating with GAI are therefore proposed:

DP3: Principle of Collaboration: When designing and developing foreign language learning applications incorporating GAI, optimizing the user interaction experience and increasing learning efficiency are vital considerations. The interface should be clean and straightforward, with clearly labeled features and a consistent layout to ensure that non-technical users can easily and efficiently access high-quality learning resources through simple input.

3.4.2 LanYu System

Based on the summarized design principles (summarized in Table 4), the LanYu system (澜语) has been designed and developed as an innovative assistive Chinese learning tool incorporating GAI to provide a more professional, richer, and engaging Chinese learning experience. The LanYu system consists of a contextual understanding module, and a contextual illustration module. And the generated contents are presented in a user interface that guarantees access to high-quality Chinese learning resources and promotes an engaging and immersive learning experience (depicted in Fig. 2).



Figure 2. User interface for the LanYu system.

The contextual understanding module processes the Chinese text or phrase entered by the user and generates contextual dialogues based on it. It fine-tunes the LLM using well-designed templates to generate two sets of clear, professional, and accurate contextual dialogues in different real-life scenarios. Each set of dialogues consists of four sentences with corresponding English translations to facilitate a complete understanding of the meaning of the dialogues. The contextual illustration module employs a T2I model to create images corresponding to the contextual dialogue, showing characters with a consistent style. Thus,

Table 4. Design Principles for leverage GAI in assisting foreign language learning

No.	Title	Design Principles
DP1	Principle of Content	Leverage GAI capabilities to create diverse, contextually rich examples aligned with user needs, enhancing understanding and motivation.
DP2	Principle of Trustworthy	Employ professional, reliable resources for training GAI, ensuring high-quality, accurate responses and trustworthiness of the application.
DP3	Principle of Collaboration	Design for optimal user interaction with a clean, straightforward interface for easy access to high-quality resources, enhancing learning efficiency.

integrating textual and visual elements helps the user vividly understand the meaning of Chinese characters and phrases and their practical applications within real-life contexts. The workflow is presented in Fig. 3.

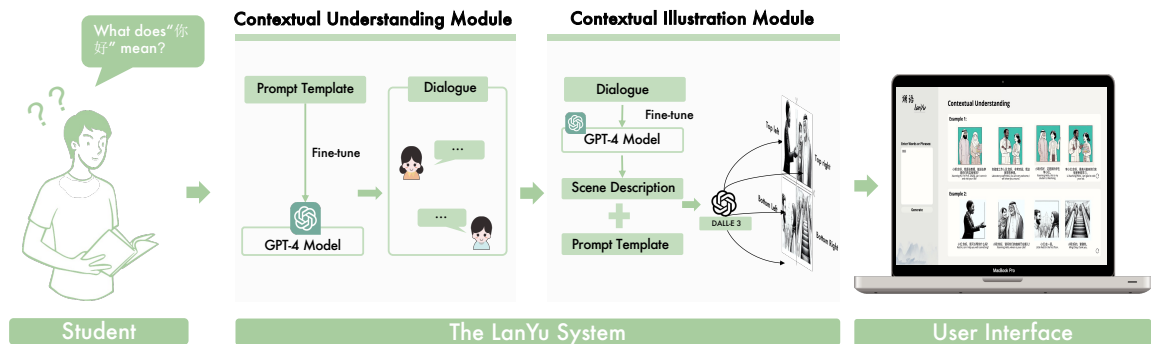


Figure 3. The workflow of LanYu system.

a) User Interface

Following the design principle DP3, the user interface of the LanYu system has been designed to be very clear and simple to ensure that users can focus on using the generated foreign language learning materials. On the left side of the user interface, there is an input box where the user can enter single or multiple Chinese characters or phrases he or she wishes to learn or understand, such as “你好” or “你好, 中文” instead of any prompts.

When the user finishes typing, he or she clicks the button below the input box, and the system transfers the input to the contextual understanding module and the contextual illustration module. These two modules transform the input into two contextual dialogues to assist the user in capturing the meaning of the input content in the specific context and generate corresponding visual illustrations so that the user can intuitively and clearly understand its application in the specific context. Not only does this allow the user to understand real-life language usage scenarios through contextual dialogues and visually

grasp the use of words and phrases in real-life contexts through accompanying images, but it also allows the user to fully concentrate on mastering and understanding language content without distraction during the learning process.

b) Contextual Understanding Module

The primary function of this module is to process Chinese characters or phrases entered by the user and convert them into dialogues of realistic contexts, thereby assisting the user in comprehending the meanings of these characters or phrases within real-world contexts. Fig. 4 illustrates the internal procedure, while Fig. 5 displays the exact code to generate contextual dialogues asynchronously.

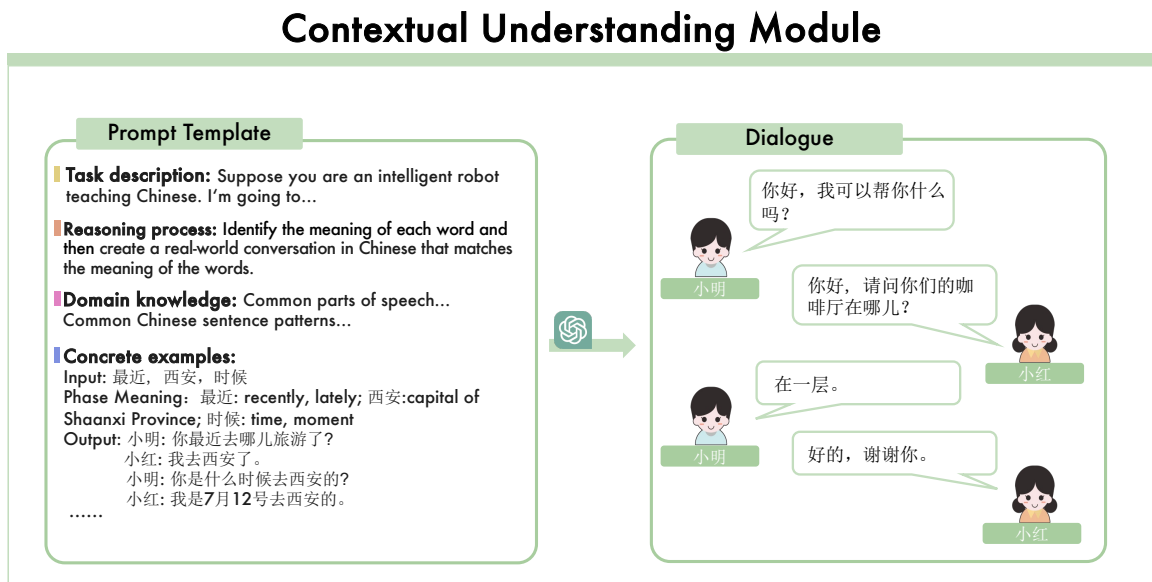


Figure 4. The internal operating process contextual understanding module.

The system utilizes a meticulously designed prompt template to fine-tune the LLM according to the design principle DP2, ensuring that the generated responses are professional and well-structured. This template guides the model in understanding and generating contextualized dialogues based on given inputs while guaranteeing that the outputs meet the requirements as supplemental instructional resources. The template is organized into four main sections: task description, reasoning process, domain knowledge, and concrete examples.

Initially, the template directs the model to concentrate on Chinese language generation by specifying the role of the LLM and clarifying its task - specifically, to generate a four-sentence question-and-answer dialogue that aligns with a realistic scenario based on the input content. Subsequently, to improve the model's reasoning capability, the template offers guidance based on CoT to support its reasoning process. This reasoning process involves first enabling the model to comprehend the lexical meanings of each Chinese character or

phrase and then constructing real-world dialogues based on these meanings. Furthermore, to ensure the grammatical accuracy of the generated sentences, the templates also contain information regarding frequently used syntax and lexical properties, which assists the model in constructing contextual dialog more accurately.

```
# Contextual understanding module to generate text based on the input values
async def async_function_segment_text(input_value1, input_value2):
    # Asynchronously creates a session for making HTTP requests
    async with aiohttp.ClientSession() as session:
        # The first contextual dialogue
        response1 = await session.post(
            url='https://api.openai.com/v1/chat/completions', # URL for the OpenAI API endpoint
            headers={'Authorization': f'Bearer {api_key2}'}, # The API key
            json={"model": "gpt-4", # Use GPT-4 model to generate contextual dialogue
                 "messages": [
                     {"role": "system", "content": prompt_seg}, # Prompt template
                     {"role": "user", "content": input_value1}, # User Input
                 ]
            })
        data1 = await response1.json()

        # The second contextual dialogue
        response2 = await session.post(
            url='https://api.openai.com/v1/chat/completions',
            headers={'Authorization': f'Bearer {api_key2}'},
            json={"model": "gpt-4",
                 "messages": [
                     {"role": "system", "content": prompt_seg},
                     {"role": "user", "content": input_value2},
                 ]
            })
        data2 = await response2.json()

    # Returns the contextual dialogue from GPT-4 model
    return data1['choices'][0]['message']['content'], data2['choices'][0]['message']['content']
```

Figure 5. The code snippet of asynchronously generated contextual dialogues.

The system is provided with 20 sets of contextual dialogues in a real-world context from the professional textbook, *New Concept Chinese 2* (Yonghua, 2013), with the structure of “*Input, Phrase Meaning, Output*” to ensure that generated contents are professional and structured. These examples allow the LLM to learn and imitate professional language usage and produce output with the same structure. Finally, the user input, along with this template, is processed through the GPT-4 model and asynchronously generates the desired two sets of contextual dialogues in different real-world scenarios.

Through this well-designed template, the LanYu system effectively assists users in mastering Chinese in a more easy-to-understand and practical way by providing high-quality educational content that is contextually relevant to real-life situations, thus improving learning

efficiency and experience.

c) Contextual Illustration Module

The LanYu system has developed the contextual illustration module in Fig. 6 based on the design principle DP1 to generate the visual representation corresponding to the generated contextual dialogues.

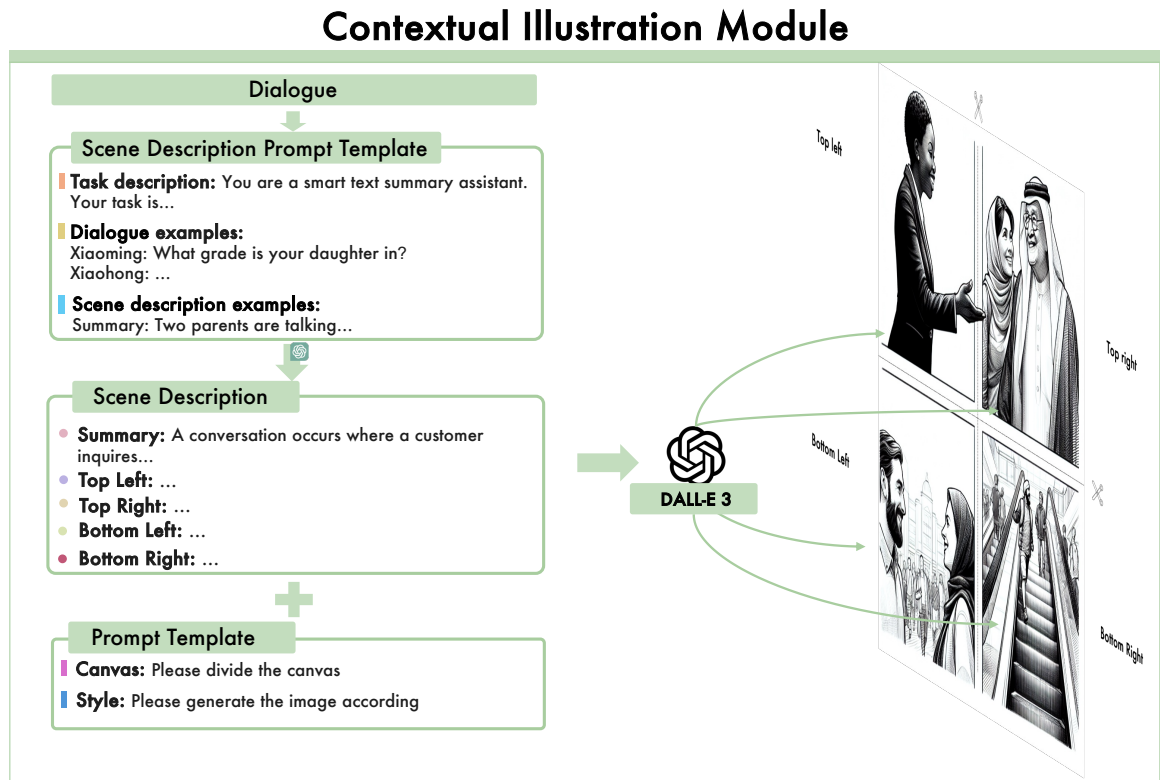


Figure 6. Contextual illustration module workflow.

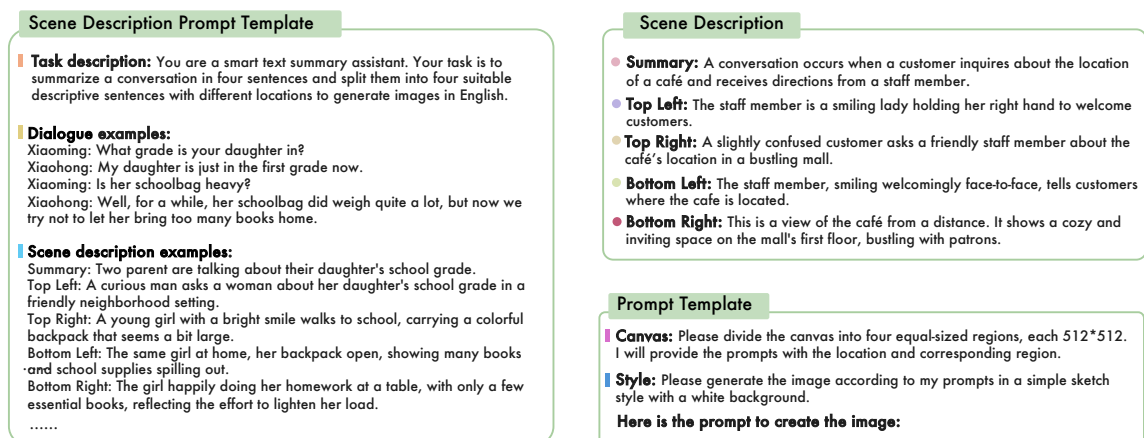


Figure 7. The prompt template of the contextual illustration module.

This module again requires fine-tuning of the LLM to accurately capture and summarize the

contextual dialogues, transforming them into concrete scene descriptions with few examples. For instance, a dialogue such as “Xiaoming: What grade is your daughter in?” would be transformed by the GPT-4 model into a scene description like “A curious man is asking a woman about her daughter’s school grade in a friendly neighborhood setting.” This detailed description of the scene and characters allows the T2I generation model to precisely capture the intended visual context and produce images that closely correspond to contextual dialogues.

```
# Asynchronously creates a session for making HTTP requests
async with aiohttp.ClientSession() as session:
    # The contextual illustration corresponding first generated contextual dialogue
    response1 = await session.post(
        url: 'https://api.openai.com/v1/images/generations',
        headers={'Authorization': f'Bearer {api_key3}'}, # The API key
        json={"model": "dall-e-3", # DALL-E 3 model
              "prompt": prompt_image + input_value1, # Prompt template and user input
              "n": 1, # generate one image
              "size": "1024x1024", # the size of image
              "response_format": "url"}
    )
    image1 = await response1.json()
    generated_image_1 = requests.get(image1["data"][0]["url"]).content

    # The contextual illustration corresponding second generated contextual dialogue
    response2 = await session.post(
        url: 'https://api.openai.com/v1/images/generations',
        headers={'Authorization': f'Bearer {api_key3}'},
        json={"model": "dall-e-3",
              "prompt": prompt_image + input_value2,
              "n": 1,
              "size": "1024x1024",
              "response_format": "url"}
    )
    image2 = await response2.json()
    generated_image_2 = requests.get(image2["data"][0]["url"]).content
```

Figure 8. The code snippet for generating corresponding visual illustrations.

The prompt template (shown in Fig. 7) then constructs a specialized canvas layout for the DALL-E 3 model to create images corresponding to each contextual dialog. Fig. 8 presents the implemented code. The canvas is divided into four equal-sized regions, each measuring 512x512 pixels and positioned at “Top Left, Top Right, Bottom Left, Bottom Right.” Each generated scene description is sequentially allocated to these four regions to ensure visual content alignment with textual content. Furthermore, to maintain the consistent style and clarity of each generated image, the prompt further instructed the DALL-E 3 model to produce images using a simple sketch style with a solid color background.

Therefore, this module ensures consistency in visual style and character coherence, allowing each illustration to visually reflect the corresponding contextual dialogues. In addition, a simple sketching style allows learners to concentrate on the content and avoid being distracted by complex visual elements. By combining contextual dialogues with intuitive and vivid visual illustrations, the LanYu system offers users professional and engaging learning materials that facilitate direct observation of vocabulary and expressions applied in real-world contexts and significantly improve learning efficiency, thereby enabling learners to acquire knowledge in a productive and enjoyable learning experience.

4 Demonstration and Evaluation

Upon successfully developing and implementing the LanYu system, the five experts and students who have previously participated in the interviews are re-invited to evaluate whether it meets the user's requirements. The design of the evaluation experiment is according to the TAM (Davis, 1989; Pal and Vanijja, 2020; Ibrahim et al., 2017), including three key assessment metrics: perceived usefulness, intention to use, and perceived ease of use. The evaluation result is displayed in Fig. 9.

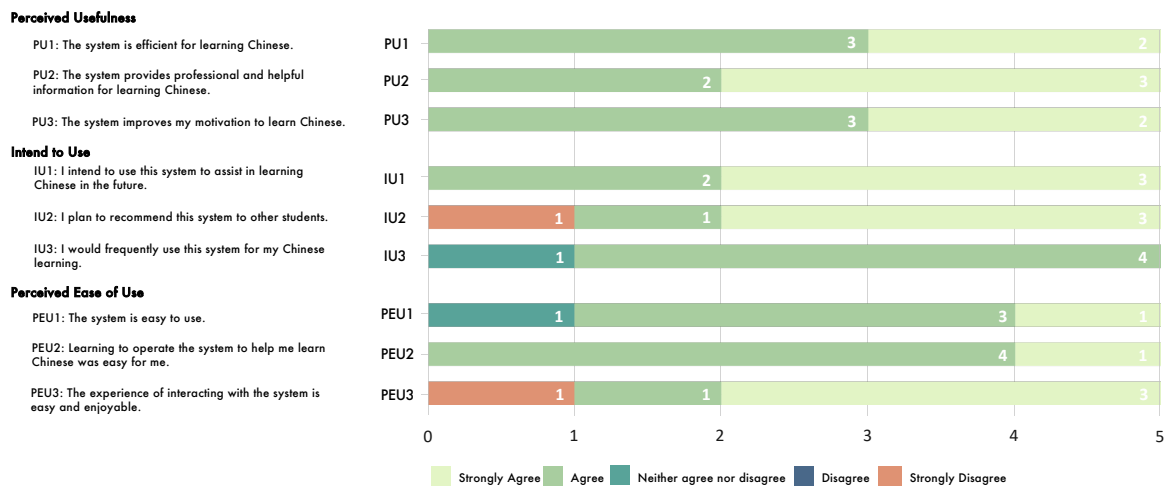


Figure 9. The results of user evaluation of the LanYu system are based on **perceived usefulness** (PU1-PU3), **intention to use** (IU1-IU3), and **perceived ease of use** (PEU1-PEU3).

Participants were initially asked to select Chinese characters or phrases from the *New Concept Chinese 2* (Yonghua, 2013) that they were interested in or unfamiliar with and enter them into the system. The system then displayed the corresponding contextual dialogues and visual illustrations based on the selections (shown in Fig. 10). Participants were free to make multiple requests until they decided on the point to end the test.

Participants were invited to evaluate the generated results and their experiences at the end of the experiment on a five-point Likert scale ranging from 1 to 5, with 1 being strongly disagree to 5 being strongly agree. The questionnaire (shown in Appendix 1) was designed through the SurveyMonkey platform (SurveyMonkey, 2024), shared with participants as a link, and took approximately 12 minutes to complete. Participants were also encouraged to comment on each system response in the questionnaire. This feedback was utilized to analyze the practical application of the system, focusing mainly on the usefulness of the interface design, the user interaction experience, and the quality of generated content. The insights gathered from this feedback helped to refine the system to better align with the needs of its users.

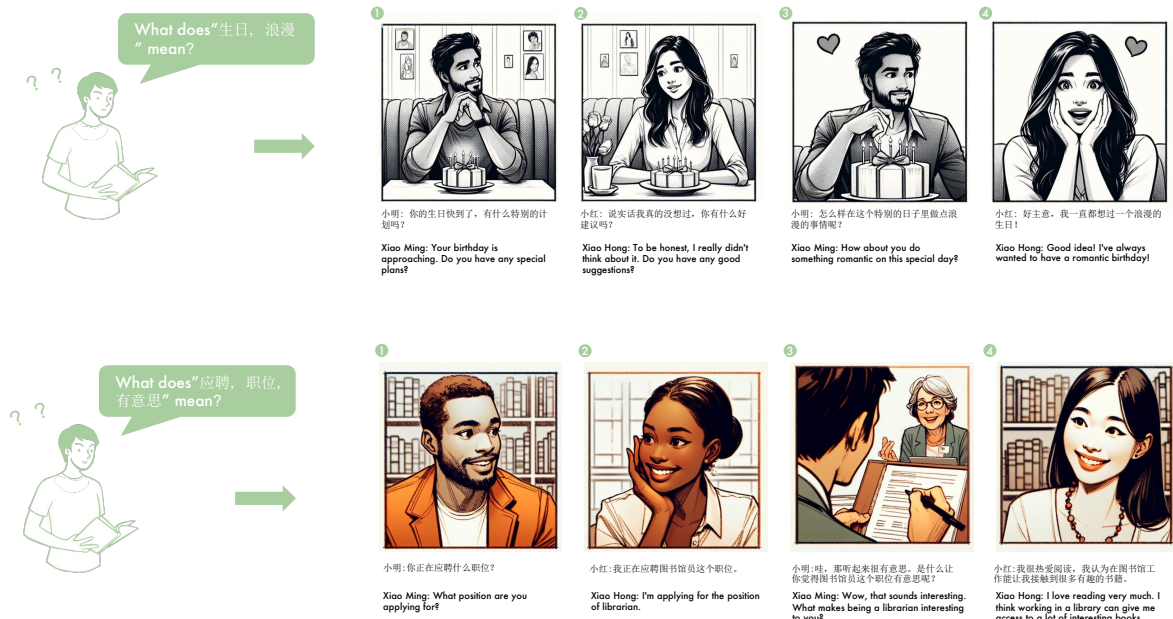


Figure 10. Some generated results when the user interacts with the LanYu system.

a) Perceived Usefulness

The general feedback from all participants is that the LanYu system effectively supported their Chinese learning and provided comprehensive, professional, and practical information.

For the generated content, participants generally agree that the information provided by the system is professional and valuable in learning Chinese. For instance, *E2* pointed out that the generated contextual dialogue was highly professional and similar to the content in the textbook. However, the only problem was that the text in the generated images sometimes confused her. This issue is further explained in the next section.

In particular, combining contextual and visual elements is incredibly beneficial for understanding and memorizing Chinese words and their meanings, improving memorization efficiency and increasing the attractiveness of the material. Participant *S1* noted that the LanYu system facilitated her learning process more effectively by integrating real-world scenarios than traditional textbook learning, enriching the learning experience and turning theoretical knowledge into practical skills in real-life situations. Similarly, *E2* emphasized the effectiveness of this system in revealing the meaning of new words in diverse contexts, and praised its skillfully combining images with vocabulary, enabling users to memorize the meaning of these words more easily. In addition, participant *S1* recommended incorporating additional elements, such as moving conversations to audio and having them spoken by native speakers, to augment the learning experience, which is discussed in detail in the next section.

Regarding motivation, all participants reported that the system greatly motivated them to

learn Chinese. *S1* highlighted learning through real-life scenario-based dialogues proves to be more engaging and inspiring compared to the traditional textbook-based learning approaches.

“I think that motivation is a quality that the student must already have, but the system makes it easier to learn new vocabulary and sentence structures, which motivates me to keep learning new things.” - E2

“I think this system is much more engaging than just using books in the classroom. I have been able to learn through real-life scenarios, and this has greatly boosted my learning.” - S1

b) Intention to Use

All participants have unanimously expressed their strong willingness to continue using this system to support their Chinese learning in the future. This novel learning method is believed to be more exciting and practical, provides a wide variety of learning materials, and dramatically expands both the horizon and the depth of knowledge. In addition, most of the participants have expressed their willingness to recommend the system to other learners, and participator *S1* is eager to actively introduce this system to anyone involved in Chinese teaching or learning. Furthermore, participants widely reported that they would frequently use the LanYu system to support their Chinese learning, reflecting a high recognition and acceptance of using this system as a practical educational resource.

“I would like to use this system to assist my learning as I have found it quite useful in giving me contextual examples of vocabulary, which is something that is often lacking in self-study materials” - E2

“When a similar system is available, I would definitely recommend it to anyone who is involved in teaching or studying Chinese.” - S1

c) Perceived Ease of Use

Most participants enthusiastically praised the high usability and user-friendliness of the LanYu system. They agreed that the experience of interacting with the system was both intuitive and enjoyable, and reported that learning how to use the system to help them learn Chinese was effortless. However, *E2* noted that she experienced some confusion when using the system for the first time because she was unfamiliar with existing similar systems integrated with GAI, indicating that the system may need to provide more guidance and support for users during the introductory phase. This issue is also discussed in more detail in the next section.

“It was very simple and engaging. Let’s put this system into real classrooms as soon as possible. This kind of new way method based on real-world scenarios makes it much easier and fun to learn.” - S1

5 Discussion and Limitation

5.1 Discussion

5.1.1 Design Principles

While GAI is rapidly becoming an essential tool for multidisciplinary applications and offers a wide range of possibilities, it also reveals the lack of clear and specific design principles in guiding the design of its practical applications. After review and iterative refinement, a recent study identified six general design principles applicable to GAI applications, as shown in Fig. 11 (Weisz et al., 2024).

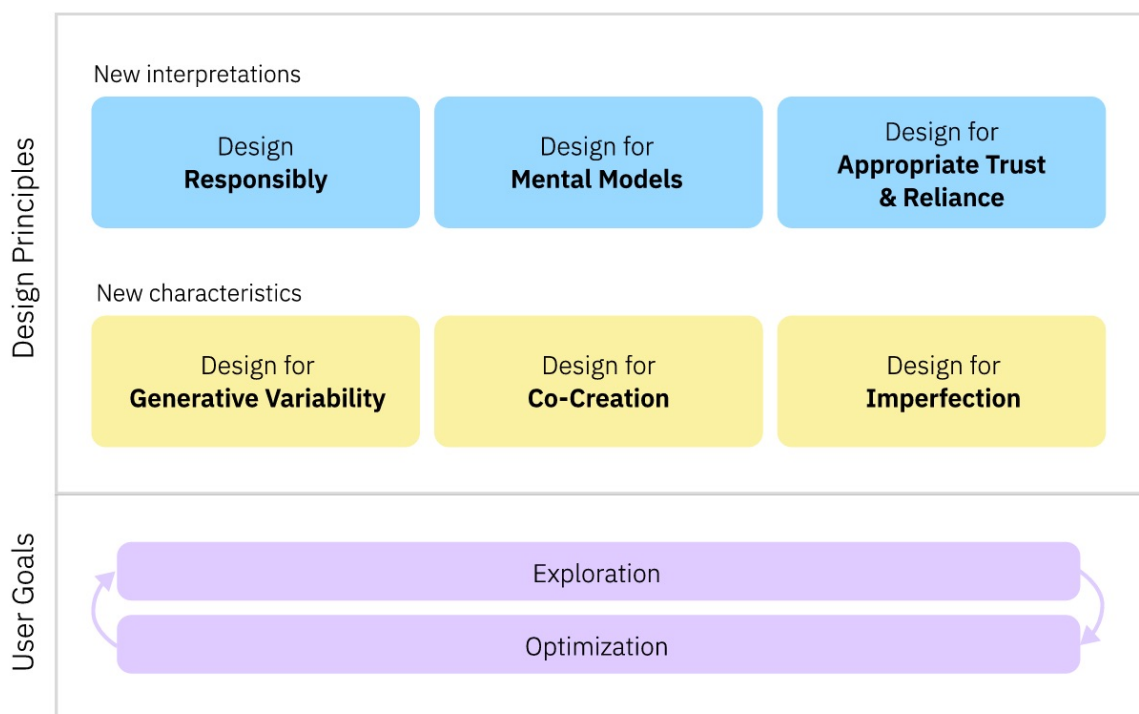


Figure 11. Six general design principles for GAI applications (Weisz et al., 2024).

Instead of exploring general GAI application design principles, this research focuses on deriving design principles for GAI applications in foreign language learning. This research is driven by a modified DSR methodology and aims to identify and define DRs through a literature review and user interviews with experts and students. Based on these requirements, three design principles specific to utilizing GAI to assist foreign language learning are derived: *DP1, the Principle of Content*; *DP2, the Principle of Trustworthiness*; and *DP3, the Principle of Collaboration*, which serve as practical, prescriptive design knowledge that provides clear guidelines for similar applications.

Design Principle DP1 encourages designers and developers to provide richer and more var-

ied content when designing future applications incorporating GAI to assist foreign language learning. User feedback suggests that this not only arouses learning interest in a foreign language but also maximizes the potential of GAI by providing a multi-sensory, immersive learning experience, thereby effectively integrating it into the learning process. Furthermore, the design principle DP2 emphasizes the significance of using professional and reliable data for developers when training the GAI. The professionalism, reliability, and accuracy of instructional materials are of the utmost importance, especially in the field of education. This principle focuses on addressing user concerns and building trust in GAI-integrated applications to better serve educational purposes. Thus, students and educators can use these technologies and applications more confidently and effectively. Based on the design principle DP3, it is recommended that complex and redundant forms of interaction should be avoided when designing GAI applications. The user interface and interaction design of applications should be intuitive and efficient so that users can easily access the high-quality content they need to fully engage in language learning.

It is worth noting that while the design principles derived in this research serve as a foundation for leveraging GAI to assist foreign language learning, they still lack comprehensive evaluation and validity from the users. The iterative process is needed to further refine and perfect these design principles so that we can unlock the potential of GAI to bring users a more efficient, interactive, and engaging learning experience.

5.1.2 LanYu System

This study has successfully designed and developed the LanYu system, a GAI-based Chinese language learning assistance system, based on summarized and derived design principles. It fine-tunes a LLM utilizing training samples collected from professional textbooks to obtain reliable contextualized dialogues in response to user queries. Furthermore, it also delivers intuitive and vivid visual displays that support learners in comprehending and memorizing Chinese characters and phrases through a series of illustrations, effectively raising their interest in learning Chinese.

Initial user evaluations have garnered approval from students and experts, demonstrating that the LanYu system is an innovative and effective tool for learning Chinese. It provides an efficient and enjoyable learning experience and complements and extends the current Chinese learning materials. In addition, the user interface design of the LanYu system has received high praise from users, especially for its intuitive, straightforward, and user-friendly layout. Integrating this system into the traditional learning process will break its limitations, resulting in learning more than just rote memorization or mechanical repetition. Therefore, implementing the LanYu system not only opens up new perspectives for educators and stu-

dents but also gives them a first-hand experience of the immense potential of GAI. This experience motivates them to more actively embrace and integrate advanced technologies into the traditional learning process, which in turn drives a revolution in education patterns.

However, while the LanYu system has exhibited strengths and potential in assisting foreign language education, participant feedback also points out that there is still room for improvement. For instance, some new users may experience confusion during initial interactions due to unfamiliarity with the system. In addition, the generated visual illustrations sometimes contain unrecognizable characters. The following section elaborates on these issues in more detail and identifies possible solutions. Future efforts aim to improve and optimize the user interface and two modules of the LanYu system to establish it as a robust learning tool that effectively supports foreign language learning, further broadens the horizon of learners, deepens the understanding of the learning material for learners, and provides an enriched learning experience.

5.1.3 Consistency in Image Generation

This research introduces a novel method for image generation in the contextual illustration module that only utilizes prompts to ensure the consistency of visual style and character features of the generated images. The valuable advantage of this approach, which distinguishes it from other image generation techniques, is that it does not require the pre-training of T2I generation models or a high demand on computational resources. These advances make it particularly suitable for resource-limited environments and non-specialist users.

The core of this approach lies in applying the textual comprehension capabilities of the LLM to the T2I generation process rather than relying on a complex image generation procedure. This method utilizes the LLM to transform the contextual conversation into scene descriptions. Furthermore, the system directly influences the final rendering of the image through precise instructions for partitioning the canvas and directing the images generated on each partition based on the corresponding scene descriptions. This approach ensures that the generated image series is consistent in style and characteristics.

Owing to its simplicity and low resource consumption, it can easily be applied to create images across a broader range of application scenarios and user groups. Users can effortlessly produce high-quality and consistent images by just inputting natural language, significantly reducing the technical barriers and expanding the potential for creative expression.

Therefore, this prompt-based image generation approach simplifies image generation processes, and promises to make image generation technology more popular and easier to use by providing a low-cost and efficient solution. This is crucial for promoting GAI-based ap-

plications in various fields, such as creative work and education, particularly for users who lack time to learn about prompt engineering or cannot afford the expensive cost of computing resources. However, the generated images sometimes contain unrecognizable text, which can affect the user experience and the clarity of the content. The coming section discusses this challenge and its potential solutions in detail.

5.2 Limitation

5.2.1 User Interface Design

While most participants commented on the clarity and easy accessibility of the user interface and the pleasant experience of interacting with the system, some feedback indicated that the interface could be confusing when used for the first time. In particular, some users were confused about what to enter into the input box without clear guidance. Therefore, future work will focus on optimizing the user interface design, such as dividing the input box into smaller sections to enable users to input Chinese characters and phrases separately and displaying examples to guide users. Inspired by the textbook, the user interface will display contextual dialogues and visual illustrations vertically and mimic the page format of the book to better align with the reading habits of the user. There will also be more interaction in the user interface, such as dragging the visual elements and allowing the user to customize the generated content.

"Initially I didn't quite understand how to use the system as I am not familiar with any existing similar systems, but it was quite easy after receiving some instruction." - E2

5.2.2 Problems with Unknown Characters in Generated Images

One significant issue when using T2I generative models is the potential presence of illegible characters in the generated visual illustrations, as presented in Fig.12. This can be caused by the model failing to thoroughly learn the associations between certain words and their visual representations during training, leading to the incorrect inclusion of these words as visual elements. Efforts have been made to simplify descriptions and reduce complex concepts in prompts to minimize the creation of confusing visual elements. However, it is still difficult to completely avoid this complex problem.

In order to mitigate this issue, image editing software (Adobe, 2024) can automatically segment text from images, which helps users to edit and remove distracting text elements while maintaining image quality. In addition, establishing a user feedback mechanism that allows users to report problems, such as illegible characters in the generated images, is required to prepare for the further identification and resolution of this problem.

By implementing these strategies, future applications can more effectively leverage the capabilities of GAI to provide users with clearer, more understandable images. Moreover, the future directions will center on reducing the likelihood of misinterpretation and confusion, improving the quality of generated images, and optimizing the user experience, ultimately making GAI more functional and widely applicable.

"Sometimes wrong words show up on the picture that it provides, which would make me a little bit confused. I hope the developers can fix this tiny problem." - S2



Figure 12. Unknown characters in generated images.

5.2.3 Delivering Multimodal Output

While interacting with the system, participants expressed the need for a more enriching learning experience by incorporating additional modes of interaction and forms of output, such as adding a voice module, which reads out dialogues with the correct tones and rich emotions. Future research will move in this direction by incorporating GAI to generate multimodal content, including audio and video, to facilitate a more productive and immersive learning experience. For example, the integration of audio output can help learners more accurately distinguish differences in pronunciation and intonation. At the same time, video content can provide more vivid and detailed representations of language use in real-world contexts. Therefore, by incorporating this multimodal learning approach, the LanYu system will provide more comprehensive information while simulating real-world communication scenarios through the visual and auditory senses, deepening memorization, and ultimately making learning a foreign language learning more engaging and efficient.

"It was pretty easy to use. I learn better when there are pictures, which makes it even easier. An additional voice option would be even better." - S1

"It would be even better if I could also listen to the conversation to hear the native pronunciation." - SI

6 Conclusions

The rapid development of GAI has brought about an unprecedented transformation in several fields, especially in education, where the application potential of GAI technology is gradually being explored. The challenges in GAI applications are particularly significant in foreign language learning. Foreign language learning requires not only the accurate mastery of grammar and vocabulary, but also the contextual application of the language, and these complexities increase the difficulty of applying GAI in this field. Currently, there is a lack of research proposing design principles that can be applied to the GAI to assist foreign language learning.

Therefore, this study adopts the DSR approach, which initiates problem identification, gathers DPs through literature review and user interview, and finally derives and implements applicable design principles. These principles are instantiated in the LanYu system, which utilizes GAI to facilitate vocabulary learning and enrich the foreign language learning experience by generating contextual dialogues with corresponding illustrations. This approach highlights, in particular, the ability of GAI to deal with the complexity and diversity of the language. It also reveals how GAI can be effectively integrated to increase the efficiency of learning and comprehension and enrich the learning experience. User evaluations of the system have shown favorable feedback, demonstrating the effectiveness and practicality of the system and offering robust support for foreign language learning.

Although the LanYu system performed nicely in improving learning efficiency and engagement, there are still shortcomings in the user interface design as well as the GAI-generated content. Therefore, future research will follow user feedback to further refine the established design principles and design a more user-friendly user interface. Potential improvements include improving the content display of the interface to make it more intuitive and easy to use, and integrating visual design elements to support more effective information delivery and user manipulation.

There is also a need for further optimization of GAI-generated content. In particular, it is vital to ensure the accuracy of the grammar and diction of the contextual dialogues it generates. This improvement is potentially achievable and uses a high-quality and extensive reliable dataset to fine-tune LLMs. In addition, generated images that convey false or misleading information must be avoided. To effectively address this issue, integrating more advanced image generation models and user feedback is required to identify and correct potentially inaccurate or misleading information. With continued refinement and improvement, the practice of leveraging GAI to assist foreign language learning will become an indispensable component of future foreign language learning processes.

References

- Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., et al. (2023). GPT-4 Technical Report. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2303.08774> [Accessed 6 Mar. 2024].
- Adobe. (2024). *Official Adobe Photoshop - Leading AI photo & design software*. Available at: <https://www.adobe.com/products/photoshop.html#> [Accessed 20 Apr. 2024].
- Athanassopoulos, S., Manoli, P., Gouvi, M., Lavidas, K. and Komis, V. (2023). The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom. *Advances in Mobile Learning Educational Research*, 3(2), pp. 818–824. doi: <https://doi.org/10.25082/AMLER.2023.02.009>.
- Avrahami, O., Hertz, A., Vinker, Y., Arar, M., Fruchter, S., Fried, O., et al. (2023). The chosen one: Consistent characters in text-to-image diffusion models. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2311.10093> [Accessed 22 Mar. 2024].
- Betker, J., Goh, G., Jing, L., Brooks, T., Wang, J., Li, L., et al. (2023). Improving image generation with better captions. *Computer Science*, 2(3). [online], p. 8. Available at: <https://cdn.openai.com/papers/dall-e-3.pdf> [Accessed 6 Mar. 2024].
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33. [online], pp. 1877–1901. Available at: https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf [Accessed 6 Mar. 2024].
- Chandra Kruse, L., Puro, S. and Seidel, S. (2022). How designers use design principles: design behaviors and application modes. *Journal of the Association for Information Systems*, 23(5), pp. 1235–1270. doi: <https://doi.org/10.17705/1jais.00759>.
- Chen, B., Zhang, Z., Langrené, N. and Zhu, S. (2023). Unleashing the potential of prompt engineering in large language models: a comprehensive review. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2310.14735> [Accessed 16 Mar. 2024].
- Collins, L. and Muñoz, C. (2016). The foreign language classroom: Current perspectives and future considerations. *The Modern Language Journal*, 100(S1), pp. 133–147. doi: <https://doi.org/10.1111/modl.12305>.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 13(3), pp. 319–340. doi: <https://doi.org/10.2307/249008>.

- Eurostat. (2023). *Foreign language learning statistics*. [online]. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_learning_statistics [Accessed 4 Mar. 2024].
- Fox, Rebecca, Corretjer, Olga and Webb, Kelley. (2019). Benefits of foreign language learning and bilingualism: An analysis of published empirical research 2012–2019. *Foreign Language Annals*, 52(4), pp. 699–726. doi: <https://doi.org/10.1111/flan.12424>.
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L. and Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27(1), pp. 70–105. doi: <https://doi.org/10.1080/09588221.2012.700315>.
- Gong, Y., Gao, X. and Lyu, B. (2020). Teaching Chinese as a second or foreign language to non-Chinese learners in mainland China (2014–2018). *Language Teaching*, 53(1), pp. 44–62. doi: <https://doi.org/10.1017/S0261444819000387>.
- Gregor, S., Chandra Kruse, L. and Seidel, S. (2020). Research perspectives: the anatomy of a design principle. *Journal of the Association for Information Systems*, 21(6), pp. 1622–1652. doi: <https://doi.org/10.17705/1jais.00649>.
- Han, L., Li, Y., Zhang, H., Milanfar, P., Metaxas, D. and Yang, F. (2023). Svdiff: Compact parameter space for diffusion fine-tuning. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2303.11305> [Accessed 22 Mar. 2024].
- Heidrich, D. and Schreiber, A. (2023). Visualizing Source Code as Comics Using Generative AI. In: *2023 IEEE Working Conference on Software Visualization (VISSOFT)*, pp. 40–44. doi: <https://doi.org/10.1109/VISSOFT60811.2023.00014>.
- Hevner, A. R., March, S. T., Park, J. and Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), pp. 75–105. doi: <https://doi.org/10.2307/25148625>.
- Ho, J., Chan, W., Saharia, C., Whang, J., Gao, R., Gritsenko, A., et al. (2022). Imagen Video: High Definition Video Generation with Diffusion Models. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2210.02303> [Accessed 6 Mar. 2024].
- Hu, B. (2010). The challenges of Chinese: A preliminary study of UK learners' perceptions of difficulty. *Language Learning Journal*, 38(1), pp. 99–118. doi: <https://doi.org/10.1080/09571731003620721>.
- Ibrahim, R., Leng, N. S., Yusoff, R. C. M., Samy, G. N., Masrom, S. and Rizman, Z. I. (2017). E-learning acceptance based on technology acceptance model (TAM). *Journal of Fundamental and Applied Sciences*, 9(4S), pp. 871–889. doi: <https://doi.org/10.4314/jfas.v9i4S.50>.

- Javaid, M., Haleem, A., Singh, R. P., Khan, S. and Khan, I. H. (2023). Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, 3(2), p. 100115. doi: <https://doi.org/10.1016/j.tbench.2023.100115>.
- JoshGreat. (2023). *8 ways to generate consistent characters (for comics, storyboards, books etc)*. [online]. Available at: www.reddit.com/r/StableDiffusion/comments/10yxz3m/8_ways_to_generate_consistent_characters_for/ [Accessed 22 Mar. 2024].
- Kaddour, J., Harris, J., Mozes, M., Bradley, H., Raileanu, R. and McHardy, R. (2023). Challenges and applications of large language models. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2307.10169> [Accessed 16 Mar. 2024].
- Kim, S., Shim, J. and Shim, J. (2023). A study on the utilization of OpenAI ChatGPT as a second language learning tool. *Journal of Multimedia Information System*, 10(1), pp. 79–88. doi: <https://doi.org/10.33851/JMIS.2023.10.1.79>.
- Lee, U., Jung, H., Jeon, Y., Sohn, Y., Hwang, W., Moon, J. and Kim, H. (2023). Few-shot is enough: exploring ChatGPT prompt engineering method for automatic question generation in english education. *Education and Information Technologies*, pp. 1–33. doi: <https://doi.org/10.1007/s10639-023-12249-8>.
- Li, J., Ren, X., Jiang, X. and Chen, C. H. (2023). Exploring the Use of ChatGPT in Chinese Language Classrooms. *International Journal of Chinese Language Teaching*, 4(3), pp. 36–55. doi: <https://doi.org/10.46451/ijclt.20230303>.
- Li, M. (2020). A systematic review of the research on Chinese character teaching and learning. *Frontiers of Education in China*, 15(1), pp. 39–72. doi: <https://doi.org/10.1007/s11516-020-0003-y>.
- Lu, Y., Bartolo, M., Moore, A., Riedel, S. and Stenetorp, P. (2022). Fantastically ordered prompts and where to find them: Overcoming few-shot prompt order sensitivity. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2104.08786> [Accessed 16 Mar. 2024].
- Mao, W. (2023). *99% character consistency with DALL-E 3*. [online]. Available at: <https://bootcamp.uxdesign.cc/99-character-consistency-with-dall-e-3-caec893fe388> [Accessed 22 Mar. 2024].
- Ministry of Education of the People’s Republic of China. (2023). *More than 25 million people in foreign countries are learning Chinese*. Available at: http://www.moe.gov.cn/fbh/live/2022/54618/mtbd/202206/t20220628_641484.html [Accessed 16 Mar. 2024].
- Möller, F., Guggenberger, T. M. and Otto, B. (2020). Towards a method for design principle development in information systems. In: *International Conference on Design Science Re-*

search in Information Systems and Technology. Springer, pp. 208–220. doi: https://doi.org/10.1007/978-3-030-64823-7_20.

Nazir, A. and Wang, Z. (2023). A comprehensive survey of ChatGPT: Advancements, applications, prospects, and challenges. *Meta-Radiology*, 1(2), p. 100022. doi: <https://doi.org/10.1016/j.metrad.2023.100022>.

Nichol, A., Dhariwal, P., Ramesh, A., Shyam, P., Mishkin, P., McGrew, B., Sutskever, I. and Chen, M. (2021). Glide: Towards photorealistic image generation and editing with text-guided diffusion models. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2112.10741> [Accessed 6 Mar. 2024].

OpenAI. (2022). *ChatGPT*. Available at: <https://chat.openai.com> [Accessed 6 Mar. 2024].

Pal, D. and Vanijja, V. (2020). Perceived usability evaluation of Microsoft Teams as an online learning platform during COVID-19 using system usability scale and technology acceptance model in India. *Children and youth services review*, 119, p. 105535. doi: <https://doi.org/10.1016/j.chilyouth.2020.105535>.

Peppers, K., Tuunanen, T., Rothenberger, M. A. and Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), pp. 45–77. doi: <https://doi.org/10.2753/MIS0742-1222240302>.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., et al. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8). [preprint]. Available at: https://d4mucfpsywv.cloudfront.net/better-language-models/language_models_are_unsupervised_multitask_learners.pdf [Accessed 6 Mar. 2024].

Ratcliffe, S. (2018). *Oxford essential quotations*. 6 ed. Oxford: Oxford University Press.

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, pp. 121–154. doi: <https://doi.org/10.1016/j.iotcps.2023.04.003>. [Accessed 23 Mar. 2024].

Renau, R. and María, L. (2016). A Review of the Traditional and Current Language Teaching Methods. 3(2). [online]. Available at: <http://hdl.handle.net/10234/162491> [Accessed 5 Mar. 2024].

Rombach, R., Blattmann, A., Lorenz, D., Esser, P. and Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10684–10695. Available at: <http://arxiv.org/abs/2112.10752> [Accessed 22 Mar. 2024].

- Ruiz, N., Li, Y., Jampani, V., Pritch, Y., Rubinstein, M. and Aberman, K. (2022). Dream-Booth: Fine tuning text-to-image diffusion models for subject-driven generation. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2208.12242> [Accessed 22 Mar. 2024].
- Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E., et al. (2022). Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2205.11487> [Accessed 6 Mar. 2024].
- Shadiev, R. and Yang, M. (2020). Review of studies on technology-enhanced language learning and teaching. *Sustainability*, 12(2), p. 524. doi: <https://doi.org/10.3390/su12020524>.
- Song, C. and Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, p. 1260843. doi: <https://doi.org/10.3389/fpsyg.2023.1260843>.
- SurveyMonkey. (2024). *SurveyMonkey: The Leading Online Survey Platform*. Available at: <https://www.surveymonkey.com> [Accessed 15 Apr. 2024].
- Tarp, S. and Nomdedeu-Rull, A. (2024). Who has the last word? Lessons from using ChatGPT to develop an AI-based Spanish writing assistant. *CIRCULO de Linguística Aplicada a la Comunicación*, 97, pp. 309–321. doi: <https://doi.org/10.5209/clac.91985>.
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M., Lacroix, T., et al. (2023). Llama 2: Open Foundation and FineTuned Chat Models. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2307.09288> [Accessed 6 Mar. 2024].
- Wang, J., Sun, Q., Chen, N., Li, X. and Gao, M. (2023). Boosting language models reasoning with chain-of-knowledge prompting. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2306.06427> [Accessed 22 Mar. 2024].
- Weisz, J. D., He, J., Muller, M., Hoefler, G., Miles, R. and Geyer, W. (2024). Design Principles for Generative AI Applications. *arXiv.org*, [preprint]. Available at: <http://arxiv.org/abs/2401.14484> [Accessed 16 Mar. 2024].
- Yin, Z., Sun, Q., Guo, Q., Wu, J., Qiu, X. and Huang, X. (2023). Do Large Language Models Know What They Don't Know? In: *Findings of the Association for Computational Linguistics: ACL 2023*. Toronto, Canada. Association for Computational Linguistics, pp. 8653–8665. doi: <https://doi.org/10.18653/v1/2023.findings-acl.551>.
- Yonghua, C. (2013). *New Concept Chinese Textbook 2*. Beijing: Beijing Language and Culture University Press. ISBN: 9787561933794.
- Young, J. C. and Shishido, M. (2023). Investigating OpenAI's ChatGPT Potentials in Generating Chatbot's Dialogue for English as a Foreign Language Learning. *International Journal of Advanced Computer Science and Applications*, 14(6). doi: <https://doi.org/10.14569/IJACSA.2023.0140607>.

Yu, H. and Guo, Y. (2023). Generative artificial intelligence empowers educational reform: current status, issues, and prospects. *Frontiers in Education*, 8, p. 1183162. doi: <https://doi.org/10.3389/educ.2023.1183162>.

Yue, Y. (2017). Teaching Chinese in K–12 schools in the United States: What are the challenges? *Foreign Language Annals*, 50(3), pp. 601–620. doi: <https://doi.org/10.1111/flan.12277>.

Żammit, J. (2023). Harnessing the power of ChatGPT for mastering the maltese language: A Journey of breaking barriers and charting new paths. In: *Machine Intelligence for Smart Applications: Opportunities and Risks*. Ed. by A. Adadi and S. Motahhir. Springer Nature Switzerland, pp. 161–178. doi: https://doi.org/10.1007/978-3-031-37454-8_8.

Zhang, G. X. and Li, L. M. (2010). Chinese language teaching in the UK: present and future. *Language Learning Journal*, 38(1), pp. 87–97. doi: <https://doi.org/10.1080/09571731003620689>.

Zhang, Z., Zhang, A., Li, M. and Smola, A. (2023). Automatic chain of thought prompting in large language models. In: *The Eleventh International Conference on Learning Representations*. [online]. Available at: <https://openreview.net/forum?id=5NTt8GFjUHkr> [Accessed 22 Mar. 2024].

Zheng, H. and Zheng, L. (2020). An Investigation on Language Programs in US Higher Institutions-A Case Study on Chinese Language Programs. *US-China Education Review B*, 10(6), pp. 257–265. doi: <https://doi.org/10.17265/2161-6248/2020.06.002>.

Appendix 1 Questionnaire

Thank you for your time and effort in participating in this questionnaire. Your input is invaluable to us.

This questionnaire aims to gather information to understand how users interact with our system and how we can improve it, thereby enhancing the system's Perceived usefulness, Intention to use, and Perceived ease of use. By participating in it, you consent to voluntary participation in this research.

All responses to the questionnaire are anonymous and will remain confidential.

* 1. What is your Tester ID number

* 2. What is your gender?

- Male
- Female
- Prefer not to say

* 3. What is your current age?

* 4. The system is efficient for learning Chinese

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

5. Feel free to explain your answer in short words

* 6. The system provides professional and helpful information for learning Chinese

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

7. Feel free to explain your answer in short words

* 8. The system improves my motivation for learning Chinese

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

9. Feel free to explain your answer in short words

* 10. I intend to use this system in assisting learning Chinese in the future

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

11. Feel free to explain your answer in short words

* 12. I plan to recommend this system to other students

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

13. Feel free to explain your answer in short words

* 14. I would frequently use this system for my Chinese learning

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

15. Feel free to explain your answer in short words

* 16. The system is easy to use

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

17. Feel free to explain your answer in short words

* 18. Learning to operate the system to help me learn Chinese was easy for me

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

19. Feel free to explain your answer in short words

* 20. The experience of interacting with the system is easy and enjoyable

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

21. Feel free to explain your answer in short words