

Incorporating Generative AI in Foreign Language Teaching Preparation: An Empirical Investigation Into the Effects of AI-Assisted Lesson Planning

Xinxin Guo¹, Hao Chen^{2*}

Department of Public Course Teaching, Nanyang Vocational College of Agriculture, Nanyang 473000, China¹
Artificial Intelligence Engineering College, Nanyang Vocational College of Agriculture, Nanyang 473000, China²

Abstract—Generative artificial intelligence, particularly large language models such as ChatGPT, has emerged as a promising educational technology tool with extensive application potential in teaching preparation. The research objectives of this study are: 1) to examine how generative AI affects foreign language teachers' teaching preparation experiences in terms of curriculum design, instructional resource development, teaching activity planning, work motivation, and efficiency; and 2) to identify the strengths, challenges, and teacher-recommended improvements associated with generative AI use in foreign language teaching preparation. This qualitative case study explored how generative AI assists foreign language teachers in their teaching preparation work, addressing a research gap regarding the benefits and challenges of this innovative approach. The study involved 13 foreign language teachers at a comprehensive university in central China who participated in a four-week training program. Qualitative data from semi-structured interviews were analyzed using thematic analysis. The findings indicate that generative AI positively affects teachers' preparation experiences in curriculum design, instructional resource development, and teaching activity planning, while enhancing motivation and efficiency. In degree-of-significance terms, the positive impact was large for curriculum design (reported by 10/13 teachers, approximately 77%), large for instructional resource development (10/13, approximately 77%), moderate-to-large for teaching activity planning (8/13, approximately 62%), large for work motivation (9/13, approximately 69%), and large for work efficiency (10/13, approximately 77%); by contrast, the impact on assessment design and teaching research was small (2/13 and 1/13, respectively). These insights contribute to understanding the utility and constraints of employing generative AI in foreign language teaching preparation and inform the development of effective teaching support strategies and training programs.

Keywords—Artificial intelligence; ChatGPT; generative AI; educational technology; teacher professional development

I. INTRODUCTION

Since its release in November 2022, ChatGPT, an advanced artificial intelligence language model developed by OpenAI, has garnered widespread attention in the educational field [1]. Built upon the Generative Pre-trained Transformer-3 (GPT-3) architecture with 175 billion parameters [2], ChatGPT generates human-like text responses [3] and demonstrates extensive natural language processing capabilities including query resolution, narrative generation, logical assessment, code

debugging, and machine translation [4]. As a general-purpose conversational agent, ChatGPT has potential applications across education, healthcare, software engineering, content creation, and business operations [5].

Existing literature reveals growing interest in generative AI's influence on educational preparation and teacher professional development. Integration of these technologies into teaching preparation demonstrates varied strategies and perceptions, showing benefits in processing pedagogical tasks [6] alongside challenges in addressing deeper instructional design issues [7]. Teacher perceptions present a nuanced picture, with research highlighting AI's potential to enhance pedagogical experience through comprehensive feedback [8], though faculty opinions on its effectiveness vary [9]. Sociocultural factors significantly influence educators' attitudes toward AI [10], emphasizing the importance of contextual considerations in AI adoption.

Generative AI integration into curriculum design and teaching methods suggests innovation potential. Natural language processing applications [11] and teacher-AI collaboration [12] indicate forward-thinking approaches to curriculum development, underscoring the need for supportive environments to maximize AI benefits in teaching preparation. AI shows promise for enhancing instructional resource development and activity design [13], with ChatGPT and prompt engineering proving effective in automatic question generation [14]. However, limitations in identifying deep pedagogical errors [7] necessitate balanced approaches that leverage AI capabilities while recognizing human expertise. Research indicates educators prioritize practical applicability of AI-generated feedback [15], and AI-based systems show utility in mapping learning outcomes and quality assurance [16].

Challenges and ethical considerations are significant to the discourse on AI in education. Concerns regarding academic integrity, the pedagogical implications of AI-generated texts [17], and the difficulties faced by educators in identifying AI-generated content highlight the complexities of incorporating AI into teaching preparation frameworks. Furthermore, the potential impact of AI on professional role devaluation among educators necessitates a strategic approach for its adoption [18]. Finally, the potential and limitations of generative tools such as ChatGPT in enhancing teaching capabilities are evident. While such technologies offer valuable applications and can improve instructional quality [19], ethical concerns and the balance

*Corresponding author.

between innovation and responsible use remain critical considerations [20]. The literature thus delineates a landscape where AI's contributions to teaching preparation are significant, yet circumscribed by challenges that require careful navigation to ensure ethical, effective, and inclusive educational practices.

Despite substantial research on AI applications in language learning, empirical studies examining how generative AI supports foreign language teachers' preparation work remain scarce. This study examines generative AI employment in foreign language teaching preparation, filling a research gap by offering insights into possibilities, advantages, and challenges of this technology. While ChatGPT's proficiency in creating summaries, quizzes, and learning materials could bolster professional development [21], concerns persist regarding reliability and ethical use of AI output in educational contexts where it may affect teaching quality and professional judgment [22]. Therefore, this study investigated generative AI's effects on teachers' preparation experiences and their views on its benefits and limitations, providing knowledge on the applicability and effectiveness of AI use in foreign language teaching preparation. The specific purpose of this study is to empirically investigate, in a single university setting in central China, how a four-week generative AI training intervention influences foreign language teachers' preparation experiences across curriculum design, instructional resource development, teaching activity planning, work motivation, and work efficiency, and to document their perceptions of the strengths, challenges, and recommendations regarding generative AI use. In contrast to prior studies that have primarily examined AI in learner-facing contexts such as student writing evaluation [15] and question generation for English education [14], this study focuses explicitly on teacher-facing preparation workflows across multiple language subjects (English, Japanese, German, French, and Spanish), an angle that has not been systematically examined in the literature.

To comprehensively investigate generative AI's role in foreign language teaching preparation, this study was guided by two research questions. First, how does generative AI implementation affect teachers' work experiences, including work efficiency, professional development, and instructional design capabilities? Second, what strengths, challenges, and recommendations emerge from teacher perspectives on generative AI use for teaching preparation? The findings

illuminate benefits and drawbacks of generative AI in foreign language teaching preparation, guiding educational administrators and researchers in developing effective teaching support strategies and training programs. A qualitative case study method was chosen because the phenomenon under investigation, namely teachers' preparation practices with an emerging technology, is process-oriented, context-dependent, and not yet well documented; case study methodology [23], [24] is therefore better suited than a large-scale survey for producing rich, contextualized descriptions of how teachers actually use generative AI and how they interpret its value and limitations in their own words. Scientifically, this study contributes early empirical evidence that can support the design of subsequent larger-scale quantitative and mixed-methods studies on AI-assisted teacher preparation, and it extends the literature on generative AI in education from student-facing applications toward teacher-facing preparation workflows. Socially, the findings can inform the design of teacher training programs and institutional guidelines that help language teachers use generative AI productively while safeguarding professional judgment, academic integrity, and student learning experiences. The remainder of this manuscript is organized as follows: Section II describes the research design, study group, program, interview protocol, data collection, and validity and reliability procedures; Section III presents the results organized by the two research questions; Section IV discusses the findings in relation to prior work, as well as pedagogical and theoretical implications, limitations, and directions for future research; and Section V concludes the paper by summarizing the novelty and implications.

II. METHODS

A. Research Design

A qualitative case study design was employed to conduct an in-depth examination of the phenomenon within its actual context and to gain insights into emerging technology applications in teaching practice [23]. Foreign language teachers from different linguistic backgrounds and teaching experience levels were considered as units of analysis. The researchers examined generative AI's effects on teachers' teaching preparation experiences and explored the tool's advantages and drawbacks. Fig. 1 illustrates the overall research design and flow.

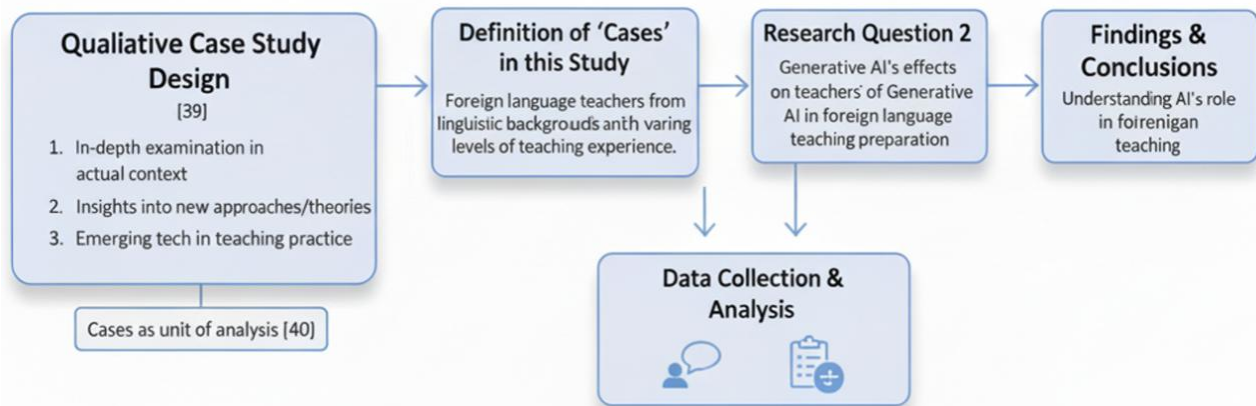


Fig. 1. Research design and flow.

B. Study Group

The study group consisted of 13 foreign language teachers at a comprehensive university in central China. Maximum variation sampling was employed to ensure diversity in linguistic backgrounds and teaching experience levels [23]. Due to administrative concerns about implementing emerging AI technology, the research was conducted with volunteers from a single college. Both training and interviews were conducted online considering teachers' schedules and geographical

dispersion. Table I presents the demographic characteristics of the participants.

As shown in Table I, the study group included 10 female and 3 male teachers aged 28-45. Teachers represented diverse linguistic backgrounds including English, Japanese, German, French, and Spanish, with teaching experience ranging from 3 to 20 years. Four participants held doctoral degrees and nine held master's degrees. Most teachers reported using ChatGPT for teaching preparation for the first time during this study.

TABLE I. DEMOGRAPHIC CHARACTERISTICS OF THE PARTICIPANTS

Teacher Code	Gender	Age	Teaching Language	Teaching Experience (years)	Education Level
T1	Female	32	English	8	Master's
T2	Female	29	English	5	Master's
T3	Female	35	Japanese	12	Doctoral
T4	Female	41	English	15	Doctoral
T5	Male	38	English	13	Master's
T6	Male	45	German	20	Doctoral
T7	Female	33	English	9	Master's
T8	Male	36	English	11	Master's
T9	Female	28	French	3	Master's
T10	Female	31	English	7	Master's
T11	Female	39	English	14	Doctoral
T12	Female	42	Spanish	18	Doctoral
T13	Female	30	English	6	Master's

C. Program Description

The program engaged teachers in four 90-minute online training sessions integrating generative AI into curriculum design, instructional resource development, teaching activity planning, and assessment design. Training modules were designed to empower teachers in enhancing their teaching preparation skills and work efficiency through the effective use of generative AI.

D. Training Materials

Training materials included tailored plans, presentation slides, and ChatGPT 3.5 free version. Curriculum design training covered unit topic design, learning objective formulation, lesson flow planning, and materials selection. Teachers used generative AI to generate instructional ideas, create teaching scenarios, and develop materials. Activity design training focused on brainstorming ideas, formulating steps, and obtaining implementation suggestions. Assessment design training involved practicing AI use for generating test items, designing scoring rubrics, and formulating formative assessment schemes.

E. Teaching Preparation Activities

Training plans comprised AI-assisted curriculum design, instructional materials development, teaching activity planning, assessment design, and teaching reflection activities. Fig. 2 shows the various ways teachers engaged with generative AI in different teaching preparation activities. As Fig. 2 illustrates, teachers used the tool to create unit lesson plans, generate learning objectives, and design teaching activities. Fig. 3 depicts the instructional design frameworks that two participants created for "Business English" and "Academic Writing" topics.

As shown in Fig. 3, teachers used the tool to generate discussion questions for group discussions and worked in pairs on AI-assisted teaching reflection exercises. Each teacher received a teaching scenario prompt to use with ChatGPT for brainstorming and planning teaching reflections, which were then shared with partners in breakout rooms for discussion. The facilitator provided standardized prompt templates initially, but teachers were encouraged to write their own prompts in the following weeks to foster independence and creativity. Teachers were guided to reference ChatGPT's answers while incorporating original professional judgment in their teaching preparation work.

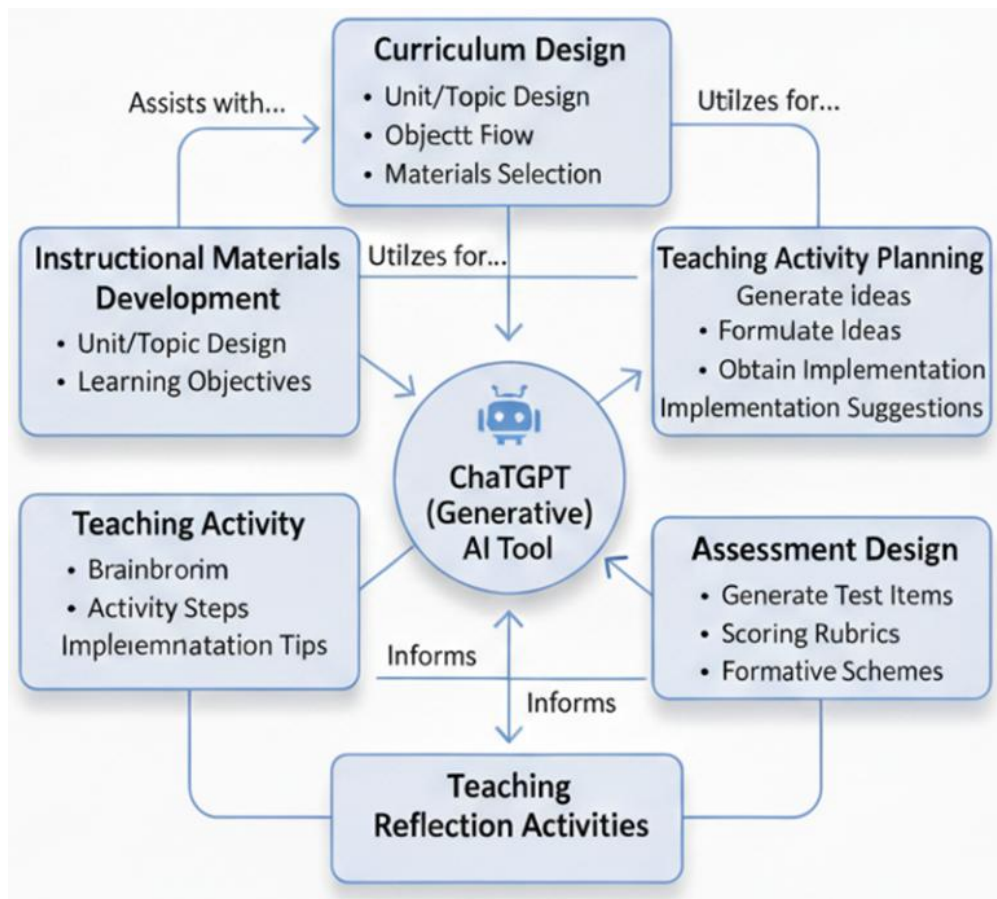


Fig. 2. Integration of ChatGPT in teaching preparation activities.

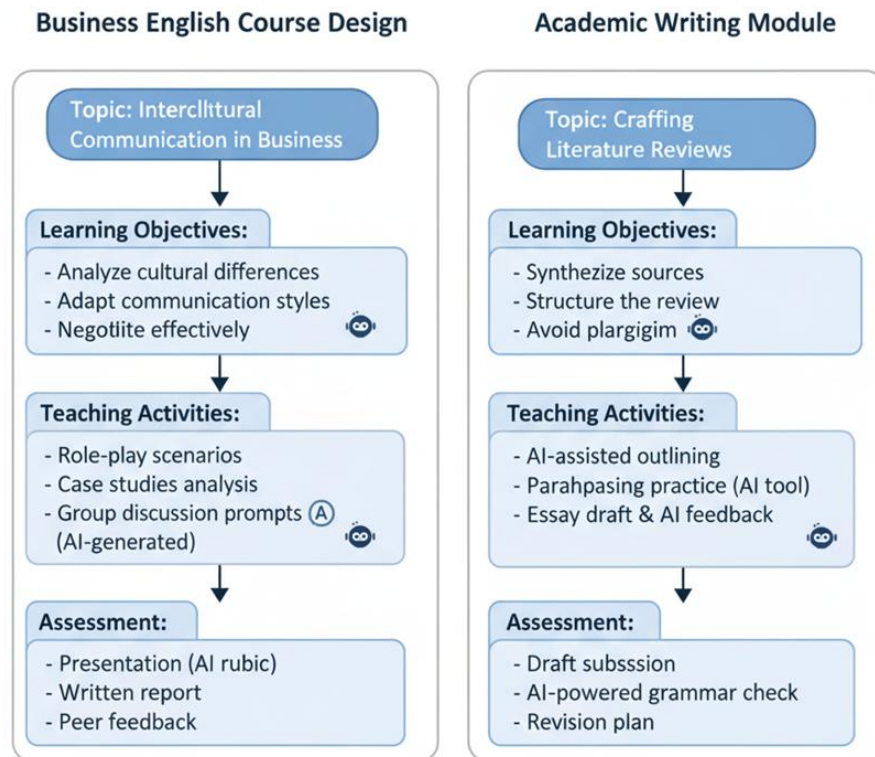


Fig. 3. Examples of instructional designs created with ChatGPT.

F. Data Collection and Analysis Process

The facilitator explained the study's purpose and data confidentiality to teachers and obtained informed consent. Teachers participated in four weeks of training activities followed by semi-structured interviews conducted online via Tencent Meeting platform. Participants indicated time preferences through a questionnaire, and interviews lasting 30-40 minutes were conducted accordingly with audio recording. The scheduling questionnaire used to obtain time preferences collected only nominal-level scheduling data (date and time-slot choices) and was used solely for logistical purposes; it was not used to generate study findings. All substantive study data were collected through open-ended semi-structured interview questions (not closed-ended survey items), and therefore no Likert-type ordinal or interval/ratio measurement scales were involved. The full interview protocol, including all open-ended questions and probes, is available from the corresponding author upon reasonable request, and a read-only version has been deposited at a public repository link that will be provided in the final version of the manuscript. Because the study's substantive instrument is a qualitative interview protocol rather than a quantitative survey, classical psychometric tests of reliability (e.g., Cronbach's alpha) and construct validity are not applicable; instead, trustworthiness was ensured through expert review of the interview guide, inter-coder reliability (simple percentage agreement of 89% and Cohen's Kappa), and maximum variation sampling (see the subsection on Validity and reliability). Given the sample size ($n = 13$, below the commonly cited $n \geq 30$ threshold for quantitative generalization), no inferential statistical tests were performed, and the findings are explicitly framed as institution-specific qualitative evidence rather than as statistically generalizable results.

Semi-structured interview questions were developed within the established case study methodology framework [24], informed by review articles on AI applications in teacher professional development [25]. Interview questions addressed generative AI's applications, impacts, benefits, drawbacks, and optimization suggestions in teaching preparation. Audio-recorded interviews were transcribed verbatim, and three researchers established a coding and category system. Each researcher independently coded the transcriptions, and coding systems were combined after reviewing similarities and differences until consensus was reached.

G. Validity and Reliability

Credibility and transferability were ensured through multiple strategies [26]. Four researchers with foreign language education experience prepared the semi-structured interview form based on training feedback. Five experts reviewed the questions for clarity, resulting in necessary amendments. The maximum variation sampling method secured transferability by examining teachers from various linguistic backgrounds and teaching experience levels. Codes and themes are clearly exhibited with direct quotations from interview data.

Reliability was enhanced through several measures [27]. Interviews were conducted immediately after training sessions to minimize contextual deviation. Researchers posed identical questions to each participant, and all interviews were recorded

for verification. Three researchers with doctoral degrees in curriculum and instruction independently coded the data. Inter-coder reliability (ICR) was calculated using the simple percentage agreement formula [28]:

$$\text{Agreement} = \frac{\text{Number of Agreements}}{\text{Total Number of Codes}} \times 100\% \quad (1)$$

Based on this formula, the three researchers achieved a simple percentage agreement of 89%, which is above the 80% threshold commonly reported as acceptable for thematic coding in qualitative research [28]. To further validate the coding consistency, Cohen's Kappa coefficient was also calculated to account for chance agreement:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (2)$$

where P_o represents the observed agreement proportion and P_e represents the expected agreement by chance. The computed Cohen's Kappa value was 0.83, indicating substantial-to-almost-perfect agreement according to the conventional interpretive ranges. It should be noted that these statistics apply to the inter-coder reliability of the qualitative thematic coding, not to a psychometric survey instrument, since no closed-ended quantitative survey was used in this study. Given that the study group ($n = 13$) falls below the commonly cited $n \geq 30$ threshold for quantitative generalization, the interview instrument is considered sufficient for qualitative case study analysis, but the findings are framed as institution-specific and are not intended as a statistically generalizable reference. As shown in Tables II and III, thematic analysis revealed emergent patterns in teacher perceptions regarding generative AI's impact on teaching preparation; the coded themes were then interpreted in alignment with the two research questions.

III. RESULTS

Thirteen teachers were interviewed to investigate generative AI-assisted teaching preparation's effectiveness on work experiences and views about its benefits and limitations. Anonymity was ensured using code names (T1, T2, T3, etc.) consistently across all tables. Teachers' views were analyzed and organized into themes, presented with detailed findings.

A. The Impact of Generative AI on Teachers' Foreign Language Teaching Preparation Experiences

The findings suggest generative AI positively impacts curriculum design, instructional materials development, and teaching activity planning (T1, T3, T4, T5, T6, T8, T10, T11, T12). Teacher T3 noted that ChatGPT helped learn new teaching strategies in authentic contexts with specific examples for classroom implementation. However, limited impact was observed on assessment design (T4, T12) and teaching research capability (T8).

Generative AI improved teachers' work motivation and professional engagement, contributing to more efficient teaching preparation experiences (T1, T7, T8, T9, T10, T12). Teacher T7 appreciated instant feedback and design suggestions that saved time previously spent consulting reference books. Participants emphasized generative AI's adaptability for interactive curriculum design, creative activity development, and exploring pedagogical approaches (T2, T3, T5, T9, T10).

Teacher T9 mentioned that ChatGPT provided cultural information about different French regions, enriching course content authenticity. Table II presents patterns of teachers'

views on generative AI's impact on teaching preparation experiences.

TABLE II. IMPACT OF GENERATIVE AI ON TEACHERS' FOREIGN LANGUAGE TEACHING PREPARATION EXPERIENCES

Themes	Category
Impact on specific teaching preparation tasks	Curriculum design capability enhancement: T5, T6, T8, T12 Instructional resource development improvement: T3, T4, T5, T10, T11, T12 Teaching activity planning advancement: T1, T8, T12 Limited impact on assessment design: T4, T12 No impact on teaching research ability: T8
Work motivation and professional engagement	Enhanced work motivation: T1, T7, T8, T9, T10, T12 Elevated professional engagement: T8, T9 No increase in work motivation: T4, T6
Versatility in teaching preparation activities	Creative teaching activity design: T3, T5 Interactive curriculum planning experiences: T2 Cross-cultural teaching content exploration opportunities: T9
Areas for improvement	Occasional technical issues: T7 Limited contribution to deep instructional design: T4

TABLE III. STRENGTHS, CHALLENGES, AND RECOMMENDATIONS ON GENERATIVE AI FOR FOREIGN LANGUAGE TEACHING PREPARATION

Themes	Category
Strengths	Instructional design capability enhancement (T1, T4, T9, T11) Time-saving and efficient resource acquisition (T1, T3, T4, T13) Comprehensive pedagogical knowledge and support (T1, T2, T10) Cross-cultural teaching content development (T1, T2, T10) Teaching vocabulary and expression expansion (T1, T5, T8, T9, T11) Pedagogical knowledge learning and examples (T3, T4, T6, T7, T8) Gamified teaching and flexible use (T1, T5)
Challenges	Over-reliance and capability deterioration concerns (T1, T2) Connection issues and AI-generated errors (T1, T3) Encouraging shortcuts in teaching preparation (T1, T2, T3) Limited effectiveness for oral teaching design (T11) Potential misuse, emotional responsiveness, and accessibility issues (T3, T5)
Recommendations for improvement	Technical enhancements and accessibility improvement (T1, T4, T5) Integration with authentic teaching contexts and interactivity (T2, T5) Novelty and engagement (T2, T5) Efficiency and time-saving capabilities (T3, T13) Addressing potential misuse and limitations (T2, T3) Improved communication and personalization (T3, T4)

Some teachers noted technical difficulties such as unstable connections or insufficiently accurate responses, and limited assistance in deep-level instructional design tasks (T4, T7). Teacher T4 observed that suggestions for complex grammar teaching sequences were often superficial, lacking consideration of students' cognitive developmental stages. Generative AI demonstrates potential to enhance teaching preparation capabilities while fostering motivation and engagement, though refinement remains essential to mitigate constraints and enhance quality.

B. Teacher Perspectives on Using Generative AI for Foreign Language Teaching Preparation: Strengths, Challenges, and Recommendations

Teachers' perspectives underscored positive impacts of generative AI in teaching preparation. Teachers praised enhanced instructional design capabilities, with T9 noting improved organization of instructional content and more creative teaching activities using diversified approaches such as task-based language teaching. Generative AI expedited teaching preparation by quickly generating instructional resources and providing pedagogical suggestions (T1, T3, T4, T13). Teacher T1 noted faster access to instructional information, with

ChatGPT providing case backgrounds, discussion questions, and role-play activities that saved preparation time.

The tool provided comprehensive pedagogical knowledge and support (T1, T2, T10) and contributed to cross-cultural teaching content development (T1, T2, T10). Teacher T10 observed that ChatGPT provided cultural information about different teaching contexts and celebration differences across English-speaking countries. Teachers mentioned improvements in teaching vocabulary and pedagogical expressions (T2, T5, T8, T9, T11) and enhanced pedagogical knowledge through practical examples (T3, T4, T6, T7, T8). Teacher T6 noted acquiring professional pedagogical terms such as "scaffolding instruction" and "differentiated instruction". Some teachers found the tool useful for gamified teaching (T1, T5). Table III presents teacher perspectives on strengths, challenges, and recommendations of generative AI in teaching preparation.

Teachers identified challenges and potential drawbacks. Concerns about over-reliance emerged (T1, T2), with T1 expressing worry about deteriorating teaching innovation capability and professional judgment. Teacher T1 noted consciously engaging in independent thinking before using ChatGPT as a verification tool rather than replacement.

Concerns included connection issues and AI-generated errors (T1, T3), with T3 recalling inaccurate Japanese honorific examples requiring careful review. Additional concerns involved encouraging shortcuts in teaching preparation (T1, T2, T3) and limited effectiveness for oral teaching design (T11). Teacher T11 noted that oral communication activity suggestions were formulaic, lacking spontaneity and interactivity.

Teachers highlighted potential misuse in assignment design (T3) and need for improved emotional responsiveness and personalized expressions (T3). Teacher T3 explained that ChatGPT's rational responses lacked humanistic warmth when addressing students' learning anxiety. Accessibility issues due to high traffic were noted (T5). Teachers offered recommendations including technical enhancements and increased accessibility (T1, T4, T5), integration with authentic teaching contexts (T2, T5), and improved communication and personalization (T3, T4). Teacher T2 suggested integrating authentic classroom videos to demonstrate strategy application effects. Teacher T4 proposed ChatGPT remember teaching style and preferences for more personalized suggestions.

IV. DISCUSSION

A. The Impact of Generative AI on Teachers' Foreign Language Teaching Preparation Experiences

Results indicate generative AI positively impacts teaching preparation capabilities in curriculum design, instructional materials development, and teaching activity planning, while increasing teacher motivation and engagement [29]. Research confirms that AI-assisted instructional design improves EFL teachers' curriculum planning, instructional organization, pedagogical coherence, and resource application compared to traditional methods [30], and stimulates authentic pedagogical interactions and productivity [31]. These findings align with our results, indicating generative AI supports teachers' professional development across multiple dimensions. The degree of change observed between traditional preparation (teachers' self-reported prior workflow) and AI-assisted preparation was substantial along several dimensions: for curriculum design and instructional resource development, 10 of 13 teachers (approximately 77%) reported a notable reduction in time required for routine tasks (e.g., generating draft lesson outlines, discussion questions, or case backgrounds) compared with their traditional practice, whereas for deep instructional design tasks (e.g., complex grammar sequencing sensitive to students' cognitive development), teachers reported only marginal benefit over traditional methods, consistent with documented limitations of ChatGPT in identifying deeper pedagogical errors [7]. The novel contribution of the present study, beyond what has been reported in [30], [31], lies in the cross-language, teacher-facing focus (English, Japanese, German, French, and Spanish) and in the explicit separation of tasks where AI substantially outperforms traditional practice (curriculum design, resource development) from tasks where it does not (assessment design, oral-communication activity design, teaching research).

Studies emphasize ChatGPT's potential for personalization and collaboration [32], complementing our observation of increased motivation and engagement. The findings are consistent with previous research supporting ChatGPT's

effectiveness in teacher professional development [33], though further investigation is needed. ChatGPT offers comprehensive feedback in instructional reflection [34] and improves instructional design capabilities through AI-generated activities and pedagogical feedback [35]. ChatGPT's potential in creating interactive dialogues and cross-cultural communication tasks [36] is particularly relevant in foreign language teaching. Multiple teachers mentioned ChatGPT's assistance in providing cultural background information and designing culturally sensitive activities, though it lacks face-to-face professional development benefits such as practical exercises in authentic contexts and emotional support.

B. Teacher Perspectives on the Advantages and Disadvantages of Generative AI in Foreign Language Teaching Preparation

ChatGPT improves instructional design capabilities by providing teaching ideas, pedagogical feedback, and introducing instructional structures and vocabulary lists, supported by studies emphasizing AI language models' contribution to teaching vocabulary skills [37]. ChatGPT fosters learner-centered teaching and enhances creativity, problem-solving abilities, and professional vocabulary development [38], with AI-assisted training showing significant improvements compared to traditional instruction [30]. ChatGPT eases preparation by saving time through efficient information provision, lesson plan generation, and materials creation [39]. ChatGPT provides opportunities for cultural and creative learning experiences, aiding critical thinking, communication skills, and intercultural competencies vital in language teaching [40], with participants noting enhanced creativity in generating teaching activities.

Participants raised concerns about possible over-reliance on ChatGPT, aligning with documented literature concerns [9]. Researchers recommend teachers maintain vital roles in teaching preparation and exercise caution in AI integration to avoid over-reliance and ensure balanced approaches incorporating technology and human insights. Concerns include emotional and psychological aspects of technology-facilitated learning [41], potential impediment to critical thinking skills and learner autonomy [42], and ChatGPT's generation of incorrect information, biases, ethical concerns, and privacy issues [43]. Teachers indicated occasional technical problems, such as connection issues [44]. These insights suggest ChatGPT's role should be viewed as complementary to traditional teaching methods, ensuring balanced approaches where teachers maximize efficiency while maintaining professional judgment and creativity.

C. Limitations and Recommendations for Future Research

The study's reliance on 13 teachers from a single institution limits generalizability across different disciplinary backgrounds, teaching experience levels, and educational environments. Future research should employ larger and more diverse sample groups across different institutional types, regions, and teacher backgrounds. Technological access and connectivity issues may have prevented full ChatGPT utilization. Teachers used the free ChatGPT 3.5 version, raising questions about whether premium versions would demonstrate different impacts. Future research could compare pedagogical effects of different ChatGPT

versions or other generative AI tools in teaching preparation. Specifically, because all participants were recruited on a volunteer basis from a single comprehensive university in central China under administrative constraints, the findings should be read as institution-specific qualitative evidence rather than as generalizable claims about generative AI in foreign language teaching broadly. In addition, the tool examined in this study was limited to ChatGPT 3.5 (free version); given the substantial capability differences between GPT-3.5 and subsequent model versions (e.g., GPT-4, GPT-4o, and later releases), the applicability of the present findings to more capable current or future generative AI tools is constrained, and the conclusions should not be read as applying to generative AI broadly. A further limitation is that the study includes no student outcome data and no objective measure of lesson quality; all reported benefits such as improved efficiency, enhanced instructional design, and increased motivation are self-reported teacher perceptions, and without external validation (e.g., classroom observation, student learning gains, or blinded expert ratings of lesson plans) it is not possible to establish whether AI-assisted preparation actually improves instructional effectiveness. Future research should therefore combine teacher self-reports with objective measures of lesson quality and student outcomes, compare model versions (e.g., GPT-4/4o and domain-specific tools), and include multi-institution samples.

Guidelines and protocols for ethical and responsible ChatGPT use in teaching preparation should be developed, clearly defining application scope, usage principles, and preventive measures for potential risks. Future research involving quantitative studies to investigate ChatGPT's effectiveness in influencing teachers' professional development outcomes, job satisfaction, teaching beliefs, and attitudes would support more valid, reliable, and generalizable findings. Adopting mixed research methods would combine qualitative depth insights with quantitative breadth coverage, providing comprehensive perspectives on generative AI's role in teaching preparation.

D. Pedagogical Implications

The study holds significant pedagogical implications for foreign language teachers. Generative AI enables teachers to personalize and manage teaching preparation processes at their own pace, unlike traditional standardized training models. AI tools provide customized suggestions based on specific needs and teaching contexts, helping teachers develop unique teaching styles and professional specializations. The interactive nature of generative AI fosters active professional engagement through curriculum design, instructional resource development, and teaching activity planning. However, ChatGPT has not yet been effective in improving oral teaching skills, whereas other AI applications like "Elsa Speak" benefit pronunciation teaching [45], suggesting the need to select appropriate AI tools for different teaching skills.

Generative AI provides varied content enabling teachers to gain insights into different cultural backgrounds and pedagogical structures, particularly important for cultivating global perspectives and cross-cultural teaching competencies. In globalization contexts, foreign language teachers must cultivate students' intercultural communicative competence, where AI tools offer unique advantages in providing multicultural

perspectives. Teachers should incorporate generative AI as supplementary support while engaging in authentic teaching practice reflection. Balancing generative AI use with professional autonomy development prevents possible teaching capability deterioration and over-reliance. Teachers should view AI as enhancing rather than replacing professional judgment, critically evaluating AI-generated content and adapting it according to specific teaching contexts while maintaining sensitivity to student needs.

E. Theoretical Implications

Generative AI integration into teaching preparation presents theoretical implications for educational theory. Generative AI demonstrates compatibility with personalized learning approaches [46], offering individualized feedback and guidance tailored to individual teaching needs. Unlike traditional "one-size-fits-all" training models, AI-supported personalized learning provides customized support based on knowledge background, teaching experience, and professional development needs, reflected in content, learning pace, and depth.

Aligned with Self-Regulated Learning theory [47], generative AI guides teachers in teaching strategy selection and time management. Teachers learn through AI interaction to set teaching preparation goals, monitor progress, evaluate outcomes, and make adjustments. This self-regulation capability cultivation is crucial for long-term professional development, enabling teachers to become active managers of professional growth. Within Cognitive Load Theory [48], generative AI's adaptable nature optimizes cognitive processing by adjusting information complexity based on teachers' knowledge level and comprehension ability, avoiding cognitive overload or resource waste. Under Engagement Theory [49], generative AI's interactive nature enhances teachers' active participation through prompt feedback and interactive learning experiences, promoting deep cognitive engagement and emotional investment that maintain sustained learning motivation and cognitive processing. These theoretical lenses are consistent with recent empirical work on ChatGPT in language education: prior IJACSA studies have demonstrated that ChatGPT can generate chatbot dialogues suitable for English-as-a-foreign-language learners at the CEFR A2 level [50] and can provide effective personalized learning experiences in higher education when guided by established instructional models such as the 5Es framework [51], offering complementary empirical support for the theoretical mechanisms discussed above.

V. CONCLUSION

In conclusion, we explored the impact of generative AI-assisted foreign language teaching preparation experiences through the views of university foreign language teachers on the opportunities and challenges of generative AI-assisted teaching preparation. The findings revealed generative AI's positive impact on curriculum design and limited impact on teaching research, development of pedagogical knowledge and instructional resource application capabilities, motivation of participants' professional development, and promotion of interactive teaching preparation experiences. Moreover, generative AI was found to contribute to participants' understanding of diverse teaching cultures and fostered their creative thinking capabilities during training activities.

However, a significant concern that emerged from the results was the potential over-reliance on generative AI. The novelty of this study lies in three aspects. First, it provides cross-language empirical evidence on teacher-facing (not student-facing) preparation workflows with generative AI across five languages (English, Japanese, German, French, and Spanish), whereas prior studies have predominantly examined AI in student-facing tasks such as writing evaluation or question generation. Second, it differentiates the magnitude of AI's impact across specific preparation sub-tasks, showing that curriculum design, instructional resource development, and teaching activity planning benefit substantially (approximately 62–77% of teachers reporting positive impact), whereas assessment design, oral-teaching design, and teaching research benefit only marginally. Third, it documents a concrete pattern of over-reliance risk that emerges even within a short four-week intervention, thereby informing institutional guidelines before large-scale adoption. The implications for science are that future research should move beyond self-reported teacher perceptions toward externally validated measures of lesson quality and student outcomes, and should systematically compare generative AI model versions. The implications for society are that universities and teacher-training programs should treat generative AI as a complement to, not a substitute for, professional judgment, and should develop ethical-use guidelines that explicitly define application scope, prevent over-reliance, and protect academic integrity. The main shortcomings of the present study, namely the small single-institution sample ($n = 13$), the exclusive use of ChatGPT 3.5, and the absence of objective lesson-quality and student-outcome measures, should be regarded as directions for further research: future work should employ larger multi-institution samples, compare more capable model versions (e.g., GPT-4/4o and successor models), and combine teacher self-reports with classroom observation and student learning data.

ACKNOWLEDGMENTS

This work was supported by the Soft Science Research Project of Henan Province (Grant No. 252400410584) and the Higher Education Research Project of Henan Province (Grant No. 2025SXHLX030).

REFERENCES

- [1] OpenAI. ChatGPT (GPT-4, September version) [Multimodal Large Language Model]. 2024. Retrieved January 02, 2024. <https://openai.com/chatgpt>
- [2] Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228-239.
- [3] Shen, Y., Heacock, L., Elias, J., Hentel, K., Reig, B., Shih, G., Moy, L. (2023). ChatGPT and other large language models are double-edged swords. *Radiology*, 307(2), 1-4.
- [4] Ouyang, L., Wu, J., Xu, J., et al. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35, 27730-27744.
- [5] Gunawan, J. (2023). Exploring the future of nursing: Insights from the ChatGPT model. *Belitung Nursing Journal*, 9(1), 1-5.
- [6] Moorhouse, B. L. (2024). Beginning and first-year language teachers' readiness for the generative AI age. *Computers and Education: Artificial Intelligence*, 6, 100201.
- [7] Algaarady, J., & Mahyoob, M. (2023). ChatGPT's capabilities in spotting and analyzing writing errors experienced by EFL learners. *Arab World English Journal*, 9, 3-17.
- [8] Bao, Y., & Li, B. (2023). A preliminary study on graduate student instructors' exploration, perception, and use of ChatGPT. *International Journal of Computer-assisted Language Learning and Teaching*, 13(1), 1-23.
- [9] Mohamed, A. M. (2024). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: Perceptions of EFL Faculty members. *Education and Information Technologies*, 29, 3195-3271.
- [10] Kim, S., & Lee, Y. (2024). Investigation into the influence of socio-cultural factors on attitudes toward artificial intelligence. *Education and Information Technologies*, 29, 9907-9935.
- [11] Bauer, E., Greisel, M., Kuznetsov, I., Bemdt, M., Kollar, I., Dresel, M., Fischer, M. R., & Fischer, F. (2023). Using natural language processing to support peer-feedback in the age of artificial intelligence: A cross-disciplinary framework and a research agenda. *British Journal of Educational Technology*, 54(5), 1222-1245.
- [12] Kim, J., Lee, H., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies*, 27(5), 6069-6104.
- [13] Ayotunde, O. O., Jamil, D. I., & Çavuş, N. (2023). The impact of artificial intelligence in foreign language learning using learning management systems: a systematic literature review. *Information Technologies and Learning Tools*, 95(3), 215-228.
- [14] Lee, U., Jung, H. C., Jeon, Y., Sohn, Y., Hwang, W., Moon, J., & Kim, H. (2024). Few-shot is enough: Exploring ChatGPT prompt engineering method for automatic question generation in English education. *Education and Information Technologies*, 29, 11483-11515.
- [15] Yang, H., Gao, C., & Shen, H. (2024). Learner interaction with, and response to, AI-programmed automated writing evaluation feedback in EFL writing: An exploratory study. *Education and Information Technologies*, 29, 3837-3858.
- [16] Zaki, N., Turaev, S., Shuaib, K., Krishnan, A., & Mohamed, E. (2023). Automating the mapping of course learning outcomes to program learning outcomes using natural language processing for accurate educational program evaluation. *Education and Information Technologies*, 28(12), 16723-16742.
- [17] Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745.
- [18] Tavares, C., Oliveira, L., Duarte, P., & Da Silva, M. M. (2023). Artificial intelligence: A blessing or a threat for language service providers in Portugal. *Informatics (Basel)*, 10(4), 81.
- [19] Alghasab, M. B. (2025). English as a foreign language (EFL) secondary school students' use of artificial intelligence (AI) tools for developing writing skills: unveiling practices and perceptions. *Cogent Education*, 12(1), 1-10.
- [20] Pack, A., & Maloney, J. (2023). Using generative artificial intelligence for language education research: Insights from using OpenAI's ChatGPT. *TESOL Quarterly*, 57(4), 1571-1582.
- [21] Khan, R. A., Jawaid, M., Khan, A. R., & Sajjad, M. (2023). ChatGPT - reshaping medical education and clinical management. *Pakistan Journal of Medical Sciences*, 39(2), 605-607.
- [22] Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, 20(2), 1-26.
- [23] Patton, M. Q. (2022). *Qualitative research and evaluation methods*. Sage.
- [24] Yin, R. K. (2011). *Applications of case study research*. Sage.
- [25] Yang, H., & Kyun, S. (2022). The current research trend of artificial intelligence in language learning: A systematic empirical literature review from an activity theory perspective. *Australasian Journal of Educational Technology*, 38(5), 180-210.
- [26] Lucas, J. W. (2003). Theory-testing, generalization, and the problem of external validity. *Sociological Theory*, 21(3), 236-253.
- [27] Cohen, L., Manion, L., & Morrison, K. (2017). *Research Methods in Education*. 8th ed. Routledge.

- [28] Miles, M., Huberman, A. M., & Saldaña, J. (2019). Qualitative data analysis: A methods sourcebook. Sage Publications.
- [29] Shadieff, R., & Liang, Q. (2023). A review of research on AR-supported language learning. *Innovation in Language Learning and Teaching*, 18(1), 78-100.
- [30] Song, C., & 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, 1260843.
- [31] Nugroho, A., Putro, N. H. P. S., & Syamsi, K. (2023). The potentials of ChatGPT for language learning: Unpacking its benefits and limitations. *Register Journal*, 16(2), 224-247.
- [32] Klayklung, P., Chocksathaporn, P., Limna, P., Kraiwanit, T., & Jangjarat, K. (2023). Revolutionizing education with ChatGPT: Enhancing learning through conversational AI. *Universal Journal of Educational Research*, 2(3), 217-225.
- [33] Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 28(11), 13943-13967.
- [34] Guo, K., & Wang, D. To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, 29, 8435-8463.
- [35] Hwang, W. Y., Nurtantyana, R., Purba, S. W. D., Hariyanti, U., Indrihapsari, Y., & Surjono, H. D. (2023). AI and Recognition Technologies to facilitate English as Foreign Language writing for supporting personalization and contextualization in authentic contexts. *Journal of Educational Computing Research*, 61(5), 1008-1035.
- [36] Topsakal, O., & Topsakal, E. (2022). Framework for a foreign language teaching software for children utilizing AR, voicebots and ChatGPT (large language models). *The Journal of Cognitive Systems*, 7(2), 33-38.
- [37] Akopiants, N. (2023). Using ChatGPT in the process of learning English: Advantages and opportunities. *Actual Problems of Ukrainian Society Development*, 1, 69-72.
- [38] Rospigliosi, P. (2023). Artificial intelligence in teaching and learning: What questions should we ask of ChatGPT? *Interactive Learning Environments*, 31(1), 1-3.
- [39] Stepanechko, O., & Kozub, L. (2023). English teachers' concerns about the ethical use of ChatGPT by university students. *Grail of Science*, 25, 297-302.
- [40] Michalon, B., & Camacho-Zúñiga, C. (2023). ChatGPT, a brand-new tool to strengthen timeless competencies. *Frontiers in Education*, 8, 1251163.
- [41] Sampson, R. J., & Yoshida, R. (2020). L2 feelings through interaction in a Japanese-English online chat exchange. *Innovation in Language Learning and Teaching*, 15(2), 131-142.
- [42] Sun, G., & Hoelscher, S. H. (2023). The CHATGPT storm and what faculty can do. *Nurse Educator*, 48(3), 119-124.
- [43] Mohammad, B., Supti, T., Alzubaidi, M., Shah, H., Alam, T., Shah, Z., & Househ, M. (2023). The Pros and Cons of Using ChatGPT in Medical Education: A Scoping Review. *Stud. Health Technol. Inform*, 305, 644-647.
- [44] Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning-are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237-257.
- [45] Akhmad, N. W., & Munawir, A. (2022). Improving the students' pronunciation ability by using Elsa Speak App. Ideas: *Journal on English Language and Learning Linguistics and Literature*, 10(1), 846-857.
- [46] Walkington, C., & Bernacki, M. L. (2020). Appraising research on personalized learning: Definitions, theoretical alignment, advancements, and future directions. *Journal of Research on Technology in Education*, 52(3), 235-252.
- [47] Zimmerman, B. J., & Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23(4), 614.
- [48] Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285.
- [49] Kearsley, G., & Shneiderman, B. (1998). Engagement theory: A framework for technology-based teaching and learning. *Educational Technology*, 38(5), 20-23.
- [50] Young, J. C., & 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): 65-72.
- [51] Albdrani, R. N., & Al-Shargabi, A. A. (2023). Investigating the effectiveness of ChatGPT for providing personalized learning experience: a case study. *International Journal of Advanced Computer Science and Applications*, 14(11):1208-1213.