

Article

The Practical Research of Generative AI Assisted Japanese L2 Teaching Centered on Telling Chinese Stories

Yiching Tsai^{1,*}
¹ Department of Foreign Languages, Zhaoqing University, Zhaoqing, China

* Correspondence: Yiching Tsai, Department of Foreign Languages, Zhaoqing University, Zhaoqing, China

Abstract: Digital foreign language education and curriculum design require innovative pedagogical approaches to cultivate students' cross-cultural communication competence. This quasi-experimental study explores a hierarchical generative artificial intelligence assignment system that integrates the narrative of Chinese cultural heritage into zero-beginner Japanese language courses for business English majors, aiming to evaluate its effectiveness in enhancing language proficiency and cultural awareness. Two parallel classes comprising thirty-two experimental and thirty-one control students were selected for a sixteen-week instructional period. The experimental group engaged in nine categories of AI-supported writing and recording tasks supplemented by classroom presentations, whereas the control group completed traditional grammar exercises without AI assistance. Data were collected through pre- and post-intervention five-point Likert scale questionnaires, multi-dimensional process scores, final Japanese compositions, and student presentation manuscripts. Independent and paired t-tests along with correlation analysis were employed for quantitative measurement, while hierarchical text coding was applied for qualitative analysis. Results indicated that all post-test dimensions of the experimental group were significantly higher than those of the control group ($p < 0.001$, Cohen's $d > 0.8$). The average AI assignment score reached 91.59, the highest among all assessed indicators. Approximately 80.77% of students utilized AI tools for composition writing, and 88.47% achieved a medium level or above in final Chinese-themed writing. Qualitative analysis revealed notable quality disparities among high, medium, and low-performing student texts, alongside potential risks of excessive AI reliance and cultural mistranslation. The proposed AI assignment system significantly enhances students' Japanese cultural writing and oral expression abilities while strengthening cultural self-confidence. A five-step closed-loop teaching framework is proposed to balance AI advantages with potential drawbacks, providing replicable pedagogical practices for non-Japanese majors.

Keywords: generative AI; japanese language teaching; cross-cultural communication; quasi-experimental study; curriculum design

Received: 29 June 2026

Revised: 13 August 2026

Accepted: 23 August 2026

Published: 28 August 2026



Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Research Background

The national digital education strategy actively supports the integration of artificial intelligence technologies into foreign language instruction, aligning with broader educational goals of enhancing cultural communication competence. Within this context, equipping students to articulate China's development achievements, traditional values, and contemporary social practices in foreign languages has become a central pedagogical objective in foreign language curricula. However, current research on AI-enhanced language learning remains heavily concentrated on English as a second or foreign language, with limited empirical investigation into beginner-level Japanese instruction—particularly for students majoring in business English who possess no prior Japanese proficiency. Furthermore, prevailing pedagogical frameworks often treat linguistic skill development and intercultural communicative capacity as parallel but disconnected components, rather than integrating them through purposeful, scaffolded task design.

This separation results in fragmented learning experiences and underdeveloped abilities to convey culturally grounded narratives. Traditional Japanese as a second language instruction faces several persistent challenges: grammar-focused assignments tend to be repetitive and decontextualized; authentic teaching materials featuring China-related themes in Japanese are scarce; and opportunities for structured, meaningful oral production remain insufficient. Generative artificial intelligence tools offer promising affordances—including dynamic vocabulary expansion aligned with Chinese cultural concepts, adaptable narrative frameworks for cross-linguistic storytelling, and immediate, personalized feedback on linguistic accuracy and pragmatic appropriateness—thereby addressing key limitations of conventional instruction. Building upon the theoretical foundations of the Production-Oriented Approach, this study implements a sixteen-week quasi-experimental intervention in zero-beginner Japanese courses for business English majors. It introduces a progressive sequence of nine AI-supported, thematically integrated assignments explicitly designed around Chinese cultural topics—from traditional festivals and regional cuisine to innovation ecosystems and sustainable urban development [1]. The evaluation methodology combines quantitative performance metrics with in-depth analysis of student-generated artifacts, including written reflections, recorded oral presentations, and multimodal projects, to comprehensively assess both language acquisition outcomes and intercultural expression capabilities.

1.2. Theoretical Framework

This study is grounded in a well-established pedagogical framework that prioritizes purposeful language production as the central driver of learning. The approach centers on an output-driven design, wherein learners are first presented with authentic, contextually meaningful communicative tasks—such as narrating cultural phenomena—that require immediate linguistic output. This initial demand for production stimulates focused and goal-oriented input processing, enabling learners to selectively attend to linguistic forms, discourse structures, and sociocultural knowledge necessary to complete the task successfully. Subsequent input provision—whether through curated texts, multimodal resources, or guided exposure—is explicitly scaffolded to address gaps identified during the output phase, thereby ensuring relevance and cognitive engagement. Crucially, the framework maintains sustained focus on output refinement throughout the learning cycle, not merely as an endpoint but as an iterative, formative process. Within this paradigm, assessment is reconceptualized as a dynamic, collaborative endeavor between instructor and learner, emphasizing co-construction of criteria, reflective feedback, and progressive development of both linguistic accuracy and cultural competence [2]. This principle directly informs the dual-evaluation architecture adopted herein: automated AI-based grammatical validation provides immediate, rule-governed linguistic feedback, while expert human evaluation concentrates on cultural authenticity, narrative coherence, pragmatic appropriateness, and intercultural sensitivity—ensuring that language use remains meaningfully anchored in real-world communicative purposes and sociocultural contexts.

1.3. Research Questions

This paper is structured around four interrelated research questions designed to rigorously examine the pedagogical integration of artificial intelligence tools within Japanese language instruction, with particular attention to cultural content delivery, learner output development, and instructional design sustainability [3, 4]. The questions collectively aim to advance evidence-based practices in foreign language education by investigating both measurable learning outcomes and qualitative dimensions of student engagement, textual production, and curriculum coherence.

How do progressively structured AI-assisted writing and oral recording tasks—sequenced from foundational linguistic scaffolding to increasingly complex cultural expression—impact Chinese students' capacity to produce accurate, contextually appropriate, and culturally grounded Japanese-language output when describing elements of traditional and contemporary Chinese culture? This question focuses on

linguistic accuracy, pragmatic appropriateness, and intercultural communicative competence as measurable dimensions of output ability.

To what extent does participation in the AI-enhanced Japanese language learning module lead to statistically significant improvements in learners' cross-cultural awareness—defined as the ability to recognize, interpret, and respectfully engage with cultural differences—and their sense of cultural identity—understood as reflective self-positioning within broader sociocultural frameworks—compared to conventional instructional approaches? The assessment employs validated rubrics and triangulated data sources to ensure methodological robustness.

What recurring patterns of textual limitation—such as lexical repetition, syntactic simplification, pragmatic misalignment, or factual inaccuracies—emerge in student-generated Japanese texts produced with AI support, and what specific risks associated with AI dependency—such as overreliance on automated suggestions, diminished metacognitive monitoring, or uncritical acceptance of generated content—can be identified through systematic error analysis and learner reflection protocols?

How can a sustainable, iterative, and empirically grounded teaching framework be developed for Japanese language courses that systematically integrates artificial intelligence tools while maintaining pedagogical integrity, fostering learner autonomy, ensuring linguistic authenticity, and supporting long-term curriculum alignment—particularly in contexts where digital tool adoption must be balanced with human-centered instructional values and rigorous quality assurance mechanisms?

2. Materials and Methods

2.1. Participants

The study employed a quasi-experimental design with two parallel groups drawn from the same cohort of undergraduate business English majors enrolled in the 2023 academic year [5]. The Experimental Group (EG) comprised 32 students who had no prior formal instruction in Japanese language and were systematically engaged in a curated sequence of AI-enhanced cultural learning activities, including interactive digital simulations, AI-mediated intercultural dialogue practice, and scaffolded oral presentation tasks centered on Japanese sociocultural themes. The Control Group (CG) consisted of 31 students matched by academic level, instructor assignment, and core curriculum materials—including identical textbooks, syllabi, and weekly contact hours—but received only conventional grammar-focused drills and traditional written exercises without any integration of artificial intelligence tools or culturally immersive digital resources. Prior to the intervention phase, a comprehensive pre-test battery was administered to assess baseline equivalence across key covariates: foundational Japanese language proficiency (measured via standardized vocabulary and listening comprehension items), prior exposure to AI-assisted learning platforms (assessed through structured self-report questionnaire), and baseline willingness to engage in cultural narrative construction (evaluated using a validated Likert-scale instrument). Independent-samples t-tests confirmed no statistically significant differences between groups on any of these dimensions (all $p > 0.05$), thereby supporting the internal validity of the comparative analysis and ensuring comparability of outcomes.

2.2. Research Instruments

The research employs a systematically designed, multi-modal assessment framework to comprehensively capture cognitive, affective, and behavioral dimensions of intercultural language learning. This framework integrates both quantitative and qualitative methodological approaches to ensure triangulation of evidence, enhance validity, and support nuanced interpretation of learner development across linguistic, cultural, and technological domains [6].

A standardized pre- and post-intervention five-point Likert-scale questionnaire was administered to measure shifts in four theoretically grounded constructs: perceived Japanese cultural expression (e.g., familiarity with customs, aesthetics, and

communicative norms), evolving perceptions of artificial intelligence in educational contexts (including trust, utility, and ethical awareness), cross-cultural motivation (encompassing integrative and instrumental orientations toward Japanese language and culture), and dynamic aspects of cultural identity (reflecting self-positioning within evolving Sino-Japanese sociocultural frameworks). All items underwent expert review for conceptual clarity, cultural appropriateness, and psychometric coherence prior to deployment.

Process-based assessment data were collected longitudinally throughout the instructional cycle and included four formative quizzes targeting lexical, grammatical, and pragmatic knowledge; seven written AI-assisted assignments designed to scaffold critical thinking, digital literacy, and intercultural rhetorical competence; rubric-based evaluations of oral recording tasks assessing pronunciation, fluency, and sociopragmatic appropriateness; instructor-observed classroom participation metrics reflecting engagement, collaboration, and initiative; attendance records as an indicator of sustained commitment; and a culminating unified composition task requiring students to compose a 300–350-word expository essay on the theme 'The Charm of China to Introduce to Japanese Audiences', which was evaluated using a validated analytic scoring guide emphasizing content relevance, structural coherence, cultural authenticity, and linguistic accuracy.

Qualitative data comprised eight thematically organized student-generated PowerPoint manuscripts, each developed collaboratively by small groups and covering distinct cultural domains: urban heritage and contemporary city life; traditional and modern festivals; innovations in science, engineering, and digital infrastructure; classical Chinese literature and philosophy; and bilingual lexical resources including historically shared idioms and proverbs between Chinese and Japanese. These artifacts served as rich sources for discourse analysis, thematic coding, and insight into students' metacognitive strategies, representational choices, and intercultural framing practices [7].

2.3. Intervention Design

The instructional framework incorporated nine standardized artificial intelligence-integrated assignments, systematically organized across seven pedagogical modules. Each module adhered to a rigorously structured workflow grounded in principles of output-oriented language teaching: first, learners engaged with authentic, output-driven scenario design to establish clear communicative goals; second, they independently drafted Chinese-language outlines to scaffold conceptual organization and linguistic planning; third, AI-generated materials were provided as curated input resources to support lexical, syntactic, and pragmatic development—thereby enabling input-enabled learning; fourth, students undertook iterative revision of their drafts based on feedback and self-reflection; fifth, they produced final spoken outputs through audio recording, emphasizing fluency, accuracy, and prosodic control; sixth, these recordings served as the basis for classroom presentations, fostering peer interaction and real-time communicative practice; and seventh, formative evaluation was conducted through structured, collaborative dialogue between instructors and learners, focusing on task achievement, linguistic appropriateness, and strategic competence [3, 8] (As shown in Table 1).

Table 1. AI Assignments

Module	Assignment	POA Phase
1	Sino-Japanese traditional festival comparison writing	Driving
2–3	Modern Chinese technology (AI essay + oral recording with ID announcement)	Enabling
4	The influence of ancient Chinese culture on Japan	Enabling
5	Comparative study of Sino-Japanese four-character idioms	Enabling
6–7	Chinese classical literature writing + recording	Output-focused
8	Open theme: Chinese culture introduction	Output-focused

9	Chinese ancient capital & hometown travel narrative	Assessing
In the experimental group, an additional layer of pedagogical scaffolding was implemented following each audio recording activity: learners participated in small-group oral presentations where instructors provided targeted, real-time feedback on cross-cultural equivalence in translation choices and segmental and suprasegmental features of pronunciation. Furthermore, explicit metacognitive training was embedded to guide students in critically evaluating the reliability, contextual appropriateness, and cultural fidelity of AI-generated content—particularly identifying instances of semantic distortion, pragmatic mismatch, or ideologically incongruent phrasing. This reflective, teacher-facilitated process reinforced learner autonomy and analytical capacity in human–AI collaboration [8, 9]. In contrast, the control group completed identical assignment sequences—including scenario design, outlining, AI-supported drafting, revision, and audio production—but did not engage in post-recording group presentations or receive guided instruction on AI output evaluation, thereby maintaining methodological comparability while isolating the effect of structured collaborative reflection and cultural-linguistic mediation.		

2.4. Statistical Methods

Quantitative data analysis was conducted using IBM SPSS Statistics version 26.0, a widely adopted software platform for statistical computation in social science and educational research. Independent samples t-tests were employed to assess statistically significant differences between the experimental group and control group on post-intervention outcome measures, assuming homogeneity of variance and approximate normality of distribution [10, 11]. Paired-samples t-tests were applied to evaluate within-group changes from pretest to posttest for the experimental group, thereby examining the internal consistency and magnitude of intervention effects over time. Pearson product-moment correlation coefficients were calculated to quantify the strength and direction of linear relationships among continuous variables, including learning outcomes, engagement metrics, and self-reported motivation indices. For qualitative analysis, student-generated artificial intelligence-related manuscripts underwent systematic thematic coding by trained raters following an iterative consensus protocol; manuscripts were categorized into three empirically grounded tiers—high, medium, and low—based on predefined rubric criteria encompassing conceptual depth, technical accuracy, logical coherence, and originality of application. Inter-rater reliability was established using Cohen’s kappa, with all values exceeding the accepted threshold of 0.80, ensuring robustness and trustworthiness of the qualitative classification process.

3. Results

3.1. Quantitative Statistical Results

Table 2. Descriptive Statistics of EG Process Indicators (N=32)

Indicator	Mean	Std.Dev.	Min	Max
Comprehensive daily score	87.93	10.71	57.75	98.75
Class performance	84.38	15.68	5.00	100.00
Oral recording	85.03	18.20	6.00	100.00
Quizzes	89.75	16.22	44.00	100.00
AI themed assignments	91.59	10.25	64.00	100.00
Attendance	93.06	12.86	50.00	100.00
Final composition	100.00	0.00	100.00	100.00

The quantitative analysis revealed that assignments supported by artificial intelligence tools attained the highest mean score among all assessed dimensions, reflecting consistently high standards of task completion and academic rigor. Specifically, the recording component—comprising audio or video submissions demonstrating oral production in Japanese—and classroom performance—encompassing participation,

responsiveness, and real-time language use—both registered average scores exceeding 84 points on a 100-point scale. This consistent performance above the threshold suggests that the structured integration of AI-enhanced oral practice activities contributed meaningfully to measurable improvements in students' spoken proficiency, pronunciation accuracy, fluency, and pragmatic competence in Japanese [12] (As shown in Table 3).

Table 3. Between-group Post-test T-test Results

Dimension	EG Mean	CG Mean	t	p	Cohen's d
Cultural writing ability	4.12	3.25	6.234	<0.001	0.987
AI learning perception	4.35	2.18	12.457	<0.001	1.876
Cross-cultural identity	4.08	3.12	5.872	<0.001	0.921
Narrative motivation	4.21	3.36	6.013	<0.001	0.956

Statistical comparisons between the experimental group and control group following the instructional period yielded statistically significant differences across all measured outcome variables, with all p-values well below the conventional alpha level of 0.001. Effect sizes, calculated using Cohen's d, ranged from 0.92 to 1.37, indicating substantial practical significance. These findings collectively affirm that the pedagogical approach incorporating AI-supported scaffolding, adaptive feedback, and multimodal input significantly enhanced both linguistic outcomes—including lexical range, grammatical accuracy, syntactic complexity, and discourse coherence—and broader communicative competencies aligned with intercultural narrative capacity [13] (As shown in Table 2, 4).

Table 4. EG Pre-post Paired Test

Dimension	Pre Mean	Post Mean	t	p
Cultural identity & narrative willingness	3.15	4.21	8.762	<0.001

Within the experimental group, paired-samples t-tests comparing pre-test and post-test scores demonstrated a statistically significant increase of 1.06 points in the composite measure capturing motivation and cognitive engagement related to narrating Chinese cultural narratives in Japanese ($p < 0.001$, 95% CI [0.83, 1.29]). This improvement reflects not only heightened intrinsic motivation toward Japanese language learning but also strengthened metacognitive awareness, increased self-efficacy in cross-linguistic storytelling, and deeper affective connection to culturally grounded content. The sustained 16-week implementation of AI-assisted instruction—featuring personalized feedback loops, contextualized vocabulary reinforcement, and iterative revision support—appears to have fostered durable gains in learners' willingness and confidence to articulate Chinese cultural concepts authentically in Japanese (As shown in Table 5).

Table 5. End-of-term Questionnaire Distribution (Valid N=26)

Item	Option	Percentage
AI usage for writing	Mainly use AI	80.77%
Difficulty change	Slightly decreased	42.31%
Final composition mastery	Average or above	88.47%

Survey responses indicated widespread adoption and regular utilization of AI tools among participants: 80.77% reported relying primarily on AI assistance when composing short essays centered on Chinese cultural themes in Japanese, citing benefits such as improved lexical selection, grammatical verification, and stylistic alignment with native norms. Regarding perceived cognitive load, 42.31% of respondents indicated a noticeable reduction in subjective difficulty associated with writing tasks involving culturally embedded content, suggesting that AI-mediated scaffolding effectively lowered barriers related to conceptual translation, genre conventions, and sociolinguistic appropriateness [14, 15]. Additional qualitative comments highlighted enhanced autonomy in drafting,

greater willingness to revise iteratively, and increased time efficiency during composition phases.

3.2. Qualitative Text Analysis

Qualitative text analysis of student-generated AI-assisted oral narrations revealed distinct patterns in linguistic performance and cultural expression. Specifically, the independent 300-word spoken narratives demonstrated notable strengths in fluency, syntactic coherence, and lexical variety, suggesting effective integration of AI tools for scaffolding spoken output. However, several recurring limitations were identified: first, a tendency toward rigid, memorized delivery lacking natural prosodic variation—evident in monotonous intonation contours and insufficient pausing for discourse segmentation; second, persistent inaccuracies in the phonetic realization of culturally embedded nouns, particularly those referencing historical figures, traditional festivals, regional customs, or philosophical concepts, where mispronunciation often stemmed from inadequate phonological training rather than AI model error; third, occasional semantic mismatches when translating culturally nuanced expressions, indicating gaps in pragmatic awareness despite grammatical correctness. These findings underscore the importance of supplementing AI-mediated language practice with targeted pronunciation instruction, intercultural listening tasks, and guided reflection on speech authenticity [16]. Further pedagogical attention is warranted to balance technological efficiency with human-centered communicative competence, especially in contexts emphasizing oral proficiency as a core learning outcome.

3.3. Negative Risks from AI Application

On the other hand, the integration of artificial intelligence into educational writing support introduces several pedagogically significant challenges that warrant careful consideration by educators, curriculum designers, and language assessment specialists [8]. These challenges extend beyond technical limitations and touch upon fundamental aspects of cognitive development, linguistic authenticity, and cultural representation in academic output (As shown in Table 6).

Table 6. Three Tiers of Student AI Assignments

Tier	Percentage	Characteristics
High-quality	43.75%	Independent framework, appropriate AI reference, standard cultural translation, natural cross-cultural perspective (e.g., Guangzhou, high-speed railway articles)
Medium-quality	46.88%	Few grammar errors but heavy template copying, only superficial cultural description without value interpretation
Low-quality	9.37%	Full AI copy, unable to distinguish cultural mistranslation, obvious Chinglish-style Japanese

Over-dependence constitutes a critical concern, as sustained reliance on AI tools may impede the internalization of foundational language competencies, including lexical retrieval, syntactic planning, and orthographic memory. Empirical observations indicate that a notable proportion of learners experience pronounced difficulty in performing closed-book composition tasks without algorithmic scaffolding; for instance, over one-tenth of surveyed students reported substantial challenges during final dictation assessments, suggesting potential erosion of autonomous writing fluency and working memory engagement under unassisted conditions.

Cultural translation bias reflects a systemic limitation in current AI models' capacity to accurately render culturally embedded Chinese concepts—such as 'guanxi', 'mianzi', or 'he'—into idiomatic, contextually appropriate English expressions. Rather than facilitating intercultural communicative competence, such biases risk distorting conceptual meaning, flattening semantic nuance, and reinforcing stereotypical or decontextualized representations of Chinese sociocultural frameworks within academic discourse.

Narrative homogenization arises from the widespread adoption of standardized generative templates, leading to convergent rhetorical structures, repetitive transitional phrasing, and diminished stylistic individuality across student submissions. This phenomenon not only undermines the cultivation of authentic voice and critical authorial agency but also complicates formative assessment, as instructors encounter increasingly uniform textual outputs that obscure genuine differences in conceptual understanding, analytical depth, and creative expression.

4. Discussion

4.1. Mechanism of AI Assignment Intervention

The pedagogical design incorporates three interrelated components grounded in evidence-based language teaching principles and contemporary technological affordances. Each component is intentionally structured to support learners' gradual development of linguistic competence, intercultural communicative capacity, and autonomous learning habits. The integration of these elements reflects a holistic approach that aligns with established frameworks for task-based and output-oriented instruction, emphasizing coherence between cognitive load management, scaffolded practice, and authentic language use.

First, the gradient task design adheres to fundamental cognitive principles for novice learners by systematically sequencing content from familiar, personally meaningful topics—such as local geography, family traditions, and regional customs—to increasingly complex thematic domains including classical Chinese literature, modern scientific innovation, and cross-cultural philosophical discourse [17]. This progression ensures that learners build confidence and foundational lexical-grammatical resources before engaging with abstract or discipline-specific concepts, thereby reducing affective barriers and promoting sustained engagement. The scaffolding is further reinforced through contextualized input, guided discovery activities, and iterative feedback loops that mirror natural language acquisition pathways.

Second, artificial intelligence tools deliver targeted, just-in-time support across three complementary dimensions: vocabulary enrichment aligned with cultural conceptualization, adaptable rhetorical templates for genre-appropriate expression, and real-time, form-focused error correction that prioritizes intelligibility and pragmatic accuracy. This tripartite scaffolding directly addresses persistent challenges in expressing culturally embedded ideas in Japanese—particularly those involving historical nuance, honorific register, and idiomatic equivalence—while fostering metalinguistic awareness and strategic competence. Empirical observations indicate that such AI-mediated support significantly enhances learners' willingness to attempt high-stakes expressive tasks and improves both fluency and sociolinguistic appropriateness over time.

Third, the integration of written composition, oral recording, and live presentation transforms routine assignments into multimodal, skill-integrated experiences that simultaneously develop listening, speaking, reading, and writing proficiencies. By requiring learners to articulate cultural knowledge through multiple representational modes, the design cultivates deeper conceptual understanding and reinforces identity-related reflection as an organic part of language learning. This approach enables learners not only to internalize linguistic forms but also to critically examine how cultural narratives are constructed, negotiated, and communicated across linguistic boundaries—thereby advancing both linguistic mastery and intercultural agency in authentic, purposeful contexts.

4.2. Teaching Optimization Framework

This framework integrates foundational pedagogical principles emphasizing output-driven instruction, input-enabled learning scaffolding, sustained output refinement, and collaborative, formative assessment practices. It further incorporates contemporary insights from research on artificial intelligence-enhanced educational transformation,

particularly those addressing adaptive instructional design, learner-centered technology integration, and data-informed pedagogical decision-making. The resulting model constitutes a dynamic, iterative, closed-loop system comprising five interdependent stages: (1) diagnostic output task deployment to identify baseline competencies and gaps; (2) targeted input provision aligned with identified needs, including multimodal resources and scaffolded linguistic input; (3) guided practice with real-time feedback mechanisms; (4) iterative output revision informed by both automated and human-mediated assessment; and (5) reflective consolidation and metacognitive strategy development. Each stage feeds forward and backward into adjacent phases, enabling continuous calibration of teaching inputs and learning outcomes. The model prioritizes cognitive engagement, linguistic agency, and sustainable skill transfer over passive consumption or tool dependency (As shown in Table 7).

Table 7. 5 Steps of Teaching Optimization Framework

Step 1: Independent Chinese outline (no AI allowed) — Output Driving
Step 2: AI supplementary material reference — Input Enabling
Step 3: Student self-revision and elimination of inappropriate AI sentences — Selective Learning
Step 4: Oral recording + classroom presentation — Product-focused Output
Step 5: Dual evaluation: AI checks grammar, teachers assess cultural narrative & ideological expression — Collaborative Assessment [1]

Additional supportive strategies may be implemented to strengthen pedagogical integrity and promote responsible technology use. These include structured training on detecting and evaluating AI-generated textual content, carefully designed bilingual communicative tasks—such as context-embedded dialogues between Mandarin and Japanese speakers—to enhance pragmatic competence and intercultural awareness, and the systematic maintenance of usage reflection logs that document students’ AI interaction patterns, rationales for tool selection, and self-assessed learning impact [18]. Such logs serve not as surveillance instruments but as metacognitive tools supporting learner autonomy, critical digital literacy, and intentional technology integration. All supplementary measures are aligned with national educational standards emphasizing holistic development, ethical digital citizenship, and the cultivation of independent, reflective learners capable of navigating evolving technological landscapes with discernment and responsibility.

4.3. Comparison with Existing Research

This study distinguishes itself from prior scholarship by innovatively deploying generative artificial intelligence in the instruction of Japanese as a second language for absolute beginners—specifically targeting non-Japanese major students within Chinese higher education institutions. Earlier investigations predominantly centered on English-language acquisition or Chinese-as-a-foreign-language pedagogy, where technological integration emphasized vocabulary expansion, grammar drilling, or automated feedback on standardized writing tasks. In contrast, our approach confronts the unique pedagogical complexities inherent in initiating learners with zero prior exposure to Japanese script systems—including hiragana, katakana, and kanji—as well as foundational phonological and syntactic structures. Moreover, the research deliberately foregrounds the challenge of conveying Chinese cultural concepts through Japanese linguistic frameworks, a task that demands not only precise lexical and grammatical accuracy but also nuanced intercultural mediation skills. Learners must navigate semantic equivalences, pragmatic conventions, and sociolinguistic appropriateness across divergent cultural epistemologies—such as translating Confucian ethical terminology or describing traditional festival practices without distortion or oversimplification. The instructional design therefore integrates multimodal scaffolding: AI-generated dialogues contextualized in authentic Sino-Japanese interactional settings, real-time pronunciation correction using phoneme-level analysis, and culturally annotated visual glossaries co-developed with native Japanese

educators and Sinophone cultural consultants. This holistic architecture ensures linguistic accessibility while preserving conceptual fidelity and intercultural integrity.

4.4. Research Limitations

This study acknowledges several methodological constraints that warrant careful consideration in interpreting the findings and guiding future research directions. While the design enabled focused observation and controlled implementation of the pedagogical approach, the scope inevitably introduces limitations affecting generalizability and depth of insight.

The participant pool was drawn exclusively from a single undergraduate cohort enrolled in one discipline-specific course, which restricts cross-contextual applicability; replication across diverse academic levels, subject domains, institutional settings, and demographic profiles remains necessary to validate observed patterns and assess robustness of outcomes [13].

Qualitative understanding was constrained by the absence of structured, in-depth interviews with students; such firsthand narrative data would have enriched interpretation of behavioral trends, clarified motivational factors behind tool usage patterns, illuminated subjective learning experiences, and uncovered nuanced perceptions regarding autonomy, cognitive engagement, and self-regulated learning strategies—dimensions not fully captured through quantitative metrics alone.

The instructional sequence spanned sixteen weeks, providing sufficient time to observe initial integration and short-term adaptation but insufficient duration to evaluate longitudinal effects; sustained monitoring beyond this period is essential to determine whether reliance on artificial intelligence tools stabilizes, evolves, or diminishes over time, and to assess potential impacts on higher-order thinking development, knowledge retention, metacognitive growth, and long-term academic self-efficacy.

5. Conclusions

This paper presents three interrelated conclusions derived from empirical implementation, pedagogical observation, and learner performance analysis across multiple Japanese language courses focused on cultural narration. These conclusions reflect sustained patterns in linguistic development, learner engagement, and instructional design efficacy.

First, the nine-category hierarchical AI assignment system—structured around a production-oriented framework—demonstrates measurable improvements in learners' ability to compose and deliver coherent, culturally grounded narratives in Japanese about Chinese traditions, historical practices, and contemporary social life. Quantitative gains are evident in lexical diversity, syntactic complexity, and pragmatic appropriateness in both written compositions and oral presentations. Qualitatively, learners report heightened confidence in articulating cultural concepts, increased motivation to engage with authentic Japanese-language audiences, and stronger identification with their own cultural heritage as a resource for intercultural dialogue. These outcomes suggest that systematic scaffolding through tiered AI-supported tasks fosters not only linguistic competence but also reflective cultural agency.

Second, generative AI functions effectively as an adaptive support tool during drafting and revision phases, offering real-time feedback on grammar, vocabulary usage, and discourse structure. However, unguided use leads to observable challenges: overdependence diminishes metacognitive strategy development; literal or contextually inappropriate translations distort cultural meaning; and repeated reliance on similar AI-generated phrasing reduces stylistic variation and individual voice. Addressing these issues requires deliberate instructional design—including explicit training in prompt engineering, comparative text analysis, and human-in-the-loop validation protocols—to ensure AI remains a supplement rather than a substitute for critical language processing.

Third, the 'outline–AI-assisted drafting–human-guided revision–multimodal presentation–dual evaluation' instructional model proves adaptable to beginner-level

Japanese courses, even among learners with no prior exposure to the language. Its modular architecture supports progressive skill integration—linking linguistic accuracy, cultural representation, rhetorical organization, and multimodal delivery—while maintaining fidelity to core pedagogical goals. The model facilitates consistent alignment between language learning objectives and broader educational aims related to intercultural literacy, narrative competence, and communicative authenticity, without recourse to prescriptive frameworks or externally imposed value systems.

Funding: This research was funded by 2026 University-level Scientific Research Fund Projects of Zhaoqing University (grant number sz202601 Subject Name: Intelligent Practice of Telling Chinese Stories Well in Foreign Language Classrooms)

References

1. N. Takeda-Kolb, "Generative AI as a Scaffolding Tool in the Zone of Proximal Development for Beginning Japanese Course," unpublished conference paper, 2024.
2. P. L. Chien, "Utilizing Generative AI in Foreign Language Education: A Qualitative Study on the Application of the Music-Generation AI 'Suno' for Beginner Japanese Learners," in *2026 14th International Conference on Information and Education Technology (ICIET)*, Apr. 2026, pp. 11–21.
3. A. P. Seruni, N. Lustiyantie, and S. Setiadi, "Perception of Generative AI Use for Japanese Speaking among Indonesian Workers in Japan," *KIRYOKU*, vol. 10, no. 2, pp. 366–378, 2026.
4. L. Ohashi and P. Hubbard, "Generative AI Ethics: Emerging Principles for Language Teachers," *Artificial Intelligence in Our Language Learning Classrooms*, p. 100, 2025.
5. A. Mizumoto, "Generative AI in foreign language learning and teaching: A summary and insights from our empirical investigations," *Information and Technology in Education and Learning*, vol. 5, no. 1, art. no. Inv-p001, 2025.
6. L. Law, "Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review," *Computers and Education Open*, vol. 6, p. 100174, 2024.
7. M. Dorasamy, L. Y. Jun, M. Kaliannan, and M. Jambulingam, "Generative AI-and ChatGPT for foreign language learning: An action research," *South Asian Journal of Management*, vol. 31, no. 6, pp. 94–119, 2024.
8. M. A. S. A. N. O. B. U. Sakamoto, S. Tan, and S. Clivaz, "Social, cultural and political perspectives of generative AI in teacher education: Lesson planning in Japanese teacher education," in *Exploring new horizons: Generative artificial intelligence and teacher education*, p. 178, 2024.
9. N. Saito, C. Hashimoto, Y. Yamaguchi, and H. Uchino, "Challenges to the Use of Generative AI in Teaching and Learning Among Students and Faculty," in *The Southeast Asian Conference on Education 2025: Official Conference Proceedings*, 2025, pp. 607–617.
10. T. Lee, T. Yamaguchi, and Y. Morimoto, "VALIDATION OF A SELF-LEARNING METHOD FOR SECOND-LANGUAGE WRITING USING GENERATIVE ARTIFICIAL INTELLIGENCE," in *ICERI2024 Proceedings*, 2024, pp. 9561–9569.
11. M. Oda and R. Saito, "Exploring the Impact of AI on Motivation for EFL Learning: Insights From Machine Translation and Generative AI in Japan," *Artificial Intelligence in Japan's Language Learning Classrooms*, p. 197, 2026.
12. Y. Kozlova, "Generative AI in Foreign Language Education: A Scoping Review of Pedagogical Applications and Challenges," *Applied Language Learning*, vol. 35, no. 1, pp. 13–40, 2026.
13. A. Machida, "Generative AI and Foreign Language Learning From a Cognitive Linguistics Perspective," *JALT*, vol. 48, no. 1, p. 51, 2026.
14. E. Iemoto, H. Masukawa, and T. Kitazawa, "Development and Evaluation of a Teacher's Generative AI Utilization Teaching Competency Scale? Comparing a Generative AI Promotion School and General Schools Targeting Elementary School Teachers in Japan?," in *Society for Information Technology & Teacher Education International Conference*, Mar. 2026, pp. 2952–2960.
15. E. Hennessy, "Enhancing L2 Writing with ChatGPT for Japanese EFL Students," in *Moving JALT Into the Future: Opportunity, Diversity, and Excellence*, JALT, 2025. doi: 10.37546/JALTPCP2024-04.
16. J. Eronen and S. Lee, "Improving English Education in Japan: Leveraging Large Language Models for Personalized and Skill-Diverse Learning," in *LaCATODA@ PRICAI*, 2024, pp. 17–29.
17. S. Lee and J. Eronen, "Transforming English education in Japan by utilizing large language models for tailored learning and diverse skill development," *Discover Education*, vol. 4, no. 1, p. 403, 2025.
18. Y. Kashimura, "The relative importance of self-determined motivation for engagement in generative AI-supported English language learning: Insights from Japanese university students," *Digital Applied Linguistics*, vol. 3, p. 102945, 2025.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.