



Generative AI in the Chinese EFL Context: A Systematic Review of Core Literature (2023–2025)

Xinyu Rui

School of Humanities and Social Sciences, Xi'an Jiaotong-Liverpool University, SuZhou,
215000, China
481180024@qq.com

Abstract. Amidst the significant pedagogical shifts triggered by Generative AI, this study conducts a hybrid systematic review of 40 pivotal studies (comprising both empirical interventions and conceptual analyses) published in core Chinese journals (CSSCI-indexed) between 2023 and 2025. The review maps the evolving landscape of Foreign Language Teaching (FLT) in China across three intersecting dimensions: the efficiency-engagement paradox, the tension between algorithmic authority and teacher autonomy, and the negotiation between global technology and local values. Findings reveal a shift in academic discourse regarding teacher identity, transitioning from “knowledge transmitters” to “human-in-the-loop collaborators.” While AI significantly enhances personalization and efficiency in exploratory settings, it introduces critical risks such as “deskilling” and cognitive loss. Navigating this gap toward a true “Human-AI Symbiosis” relies on the targeted integration of Critical AI Literacy into teacher education and the deployment of localized technical frameworks, as evidenced by the synthesized literature.

Keywords: Artificial Intelligence; FLT; Teacher Role Restructuring; Generative AI; Educational Ecosystem; Institutional Inertia

1 Introduction

The rapid development of Generative AI (GenAI) has introduced specific challenges to the educational landscape. However, the integration of GenAI into Foreign Language Teaching (FLT) is not a context-free technological transplant; it is a complex negotiation between global algorithms and local educational ecologies. In the Chinese context, this negotiation is particularly acute and presents unique theoretical value for the international dialogue on AI in education.

Unlike the typical Western classroom settings often assumed in GenAI training data and pedagogical models, Chinese FLT is characterized by distinct structural and ideological constraints. Specifically, the widespread phenomenon of “mega-sized classes” (often exceeding 60 students) challenges the standard “personalization” narrative of AI, questioning whether algorithmic efficiency can truly replace human interaction in high-density environments.

© The Author(s) 2026

C. F. Peng et al. (eds.), *Proceedings of the 2026 5th International Conference on Humanities, Wisdom Education and Service Management (HWESM 2026)*, Advances in Social Science, Education and Humanities Research 1024,

https://doi.org/10.2991/978-2-38476-593-5_18

More critically, the introduction of GenAI introduces tensions regarding knowledge construction within the Chinese framework. As GenAI models are predominantly trained on data from “WEIRD” (Western, Educated, Industrialized, Rich, and Democratic) societies, their generated content inevitably carries implicit Western values and epistemologies.

1.1 Construct Precision: Differentiating AWE and GenAI

Before analyzing the efficacy of these tools, it is theoretically imperative to delineate the construct precision between traditional Automated Writing Evaluation (AWE) systems and contemporary Generative AI (GenAI). Traditional AWE platforms operate primarily on rule-based engines, natural language processing algorithms, and shallow machine learning. Their pedagogical impact is largely confined to providing "corrective feedback" (CF) focused on lexical and grammatical accuracy. In contrast, GenAI tools are driven by Large Language Models (LLMs) and deep neural networks (e.g., Transformer architecture). GenAI is capable of generating probabilistic, context-aware text, restructuring entire paragraphs, and offering multi-dimensional feedback spanning content, organization, and pragmatics. Recognizing this paradigm shift is crucial; while foundational AWE research primarily demonstrates efficiency in mechanical error correction, the integration of GenAI fundamentally alters the pedagogical landscape by substituting higher-order cognitive design and interactive negotiation evaluation.[1]

1.2 Theoretical Lens: International Frameworks

To systematically analyze the structural tensions within the Chinese EFL context, this review adopts two cutting-edge international theoretical frameworks as its analytical lens. First, it employs the "Efficiency-Engagement Paradox" formalized by Boumediene et al..[1] Originally conceptualized in graduate dissertation writing, this paradox serves as a vital instrument to interrogate whether the rapid linguistic fluency afforded by AI in Chinese mega-sized classes comes at the severe cost of deep cognitive engagement. Second, to address the profound threat of pedagogical deskilling, the study utilizes the "Critical AI Literacy (CAIL)" framework and the associated "open box pedagogy" established by Pérez-Paredes, Curry, and Ordoñana-Guillamón.[2] These international frameworks are not conflated with the local dataset; rather, they provide the theoretical scaffolding to interpret how the 40 synthesized Chinese CSSCI publications respond to, localize, or challenge global AI dynamics.

Therefore, the core problem this review addresses is not merely how GenAI tools are used in Chinese classrooms, but how the Chinese FLT ecosystem adapts or contextually integrates these global technologies to align with its unique policy mandates.

2 Methodology

2.1 Data Sources and Retrieval Strategies

This study utilized the China National Knowledge Infrastructure (CNKI) as the primary database, strictly restricting the core qualitative scope to CSSCI-indexed journals to ensure the validity and reliability of the research in the local context.

The search keywords included “Foreign Language Teaching” and “Artificial Intelligence/Generative AI/ChatGPT,” with the time frame set from January 1, 2023, to December 31, 2025.

2.2 Inclusion and Exclusion Criteria (PRISMA Workflow)

The initial search yielded 292 results. To ensure the depth and pertinence of the review, a screening process was conducted following a strict PRISMA workflow, illustrated in Figure 1. Recognizing the multifaceted nature of educational technology integration, this review adopts a hybrid systematic approach.

Figure 1: PRISMA Flow Diagram for Systematic Review

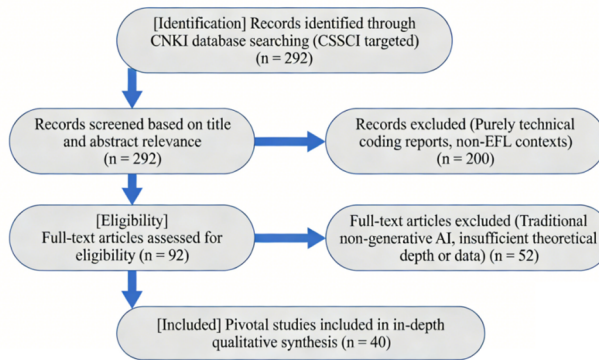


Fig. 1. PRISMA Flow Diagram for Systematic Review.

(Note: The final dataset explicitly includes both highly cited empirical interventions and foundational theoretical/conceptual analyses to map the macro-pedagogical discourse).

To provide transparency regarding the selected dataset, Table 1 summarizes the core characteristics of representative high-impact studies included in the final qualitative synthesis.

Table 1. Summary of Core Characteristics of Representative Included Literature

Author(s)&Year	Source Journal	Research Method	Sample Context	Core Findings
Xu,Hu,&Su(2024)	<i>Foreign Language World</i>	Questionnaire&Interviews	EFL University Students	AI significantly enhances writing efficiency but presents risks of passive cognitive engagement and behavioral reliance.

Kong&Yang(2025)	<i>Foreign Language World</i>	Empirical Intervention	University English Course	Proposes explicit value guides and comparative pedagogy to filter implicit Western cultural biases in LLMs like DeepSeek.
Gao(2025)	<i>Contemporary Foreign Languages Studies</i>	Mixed-Methods	Foreign Language Education	Identifies the "double-edged sword" effect of DeepSeek; excels in syntax generation but risks eroding students' independent critical thinking.
Qin(2023)	<i>Technology Enhanced Foreign Language Education</i>	Conceptual Analysis	EFL Contexts	Highlights the indispensable need for human-machine symbiosis and meta-cognitive management to prevent cognitive atrophy.
Zhang&Hong(2023)	<i>Foreign Language World</i>	Theoretical&Critical Analysis	EFL Teaching Ecosystem	Argues that while ChatGPT seemingly empowers teachers, it profoundly threatens intrinsic instructional design capacities.

Note: This table selects 5 highly influential representative papers. The selection criteria are core journals in the CSSCI, typicality of research methods, and citation frequency ranking in the top 10%.

3 Data Analysis

3.1 Bibliometric Coding

To map the basic landscape, the full dataset of 40 included publications was coded for descriptive metadata, including publication year, journal impact factor, authorship, and research methods.

3.2 Qualitative Thematic Analysis

To distill the core pedagogical tensions, the subset of 40 high-impact studies underwent a rigorous inductive thematic analysis assisted by NVivo 12 qualitative data analysis software.

The coding process followed a three-stage logic: Open Coding (Data Extraction), Axial Coding (Categorization), and Selective Coding (Theme Generation).

3.3 Research Themes and NVivo Coding Framework

Rather than applying mutually exclusive categorizations, the thematic analysis identified that the integration of GenAI is characterized by three fundamental, intersecting structural tensions. Table 2 illustrates the distribution of semantic codes and demonstrates how the three thematic tensions emerged directly from the qualitative data nodes.

Table 2. NVivo Coding Framework and Thematic Distribution Matrix

Strategy Category	Specific Forms	Example Open Coding Nodes	Number of Codes	Percentage
The Efficiency-Engagement Paradox	The Allure of Efficiency	"Instant feedback generation," "Relieving grading burden," "Automated syntax optimization"	36	90%
	The Risk of Cognitive Loss	"Blind acceptance of AI output," "Reduction in cognitive struggle," "Passive obedience"	30	75%
Algorithm Authority vs. Teacher Autonomy	The Threat of Deskilling	"Loss of curriculum design ability," "Outsourcing pedagogical evaluation"	28	70%
	Loss of Meta-Cognition	"Erosion of reflective capacity," "Inability to dynamically assess student needs"	22	55%
Global Technology vs. Local Values	Defensive Response: Epistemological Filtering	"Western logic embedded in models," "Cultural representation alienation"	18	45%
	Proactive Response: AI-Empowered Construction	"Telling China's story," "Comparative critical analysis of AI texts"	12	30%

Note: 'Number of Codes' refers to the number of publications related to this topic, and the percentage represents the relative proportion, since a single publication may cover multiple subtopics, there may be overlapping coding, so the sum of the percentages can exceed 100%.

4 Key Research Findings: Tensions and Negotiations

4.1 Theme 1: Localizing the Efficiency-Engagement Paradox

Viewed through the theoretical lens established by Boumediene et al. [1], this review finds that the "Efficiency-Engagement Paradox" is strongly confirmed and dramatically amplified within the AI intervention data of Chinese mega-sized EFL classrooms. The major tension lies between the administrative efficiency of AI and the cognitive depth of student learning.

The Allure of Efficiency in High-Density Contexts: In the Chinese EFL context, the sheer volume of student output creates a severe "feedback bottleneck" for instructors. Empirical studies within the CSSCI dataset confirm that AI tools serve as a critical release valve. For example, Xu, Hu, and Su demonstrated in their investigation of learners' behaviors in AI-assisted academic English writing that GenAI intervention positively enhances writing efficiency and structural language optimization.[3] As Wen and Liang postulate in their conceptualization of "human-AI interactive negotiation competence," GenAI transcends mechanical error correction, offering dynamic, dialogic feedback that profoundly eases the structural burden on educators.[4]

Deepening the Logic of the Local Context (The Risk of Cognitive Offloading): However, efficiency does not automatically translate to effective learning, particularly

within the ecosystem of China's "mega-sized classes." From an educational psychology and classroom ecology perspective, genuine language acquisition requires both "cognitive friction" and interpersonal emotional scaffolding. In severely high-density environments (often 60+ students) where individualized teacher-student interaction is mathematically impoverished, the sudden influx of hyper-fluent, instantaneous AI feedback acts as an "authoritative surrogate," rapidly filling the interactional void. Without the emotional resonance, motivational regulation, and pedagogical monitoring inherent in human supervision, this feedback lacks productive friction. Consequently, learners in these large cohorts are significantly more susceptible to rapid "cognitive offloading"—outsourcing the intellectual struggle of writing and revising entirely to the AI. As Gao points out regarding the "double-edged sword effect" of models like DeepSeek, overreliance bypasses the zone of proximal development, resulting in students exhibiting passive obedience to algorithmic suggestions rather than active pedagogical uptake.[5]

4.2 Theme 2: Algorithm Authority vs. Teacher Autonomy

The second prominent theme in the literature highlights severe concerns regarding professional agency, starkly juxtaposing the superficial narrative of technological "empowerment" against the insidious reality of professional "deskilling" driven by continuous cognitive offloading.

The Threat of Deskilling and the Gatekeeper Paradox: Zhang and Hong astutely observe in *Foreign Language World* that while natural language processing systems can seemingly empower foreign language teaching by rapidly generating instructional ideas, they simultaneously pose severe threats to teachers' intrinsic instructional design capacities if relied upon excessively.[6] The continuous delegation of curriculum structuring and evaluation to black-box algorithms relentlessly erodes the teacher's meta-cognitive ability. Within the specific parameters of the Chinese educational ecosystem, this dynamic generates a severe logical paradox: the "Gatekeeper-Deskilling Paradox." Current national educational mandates strictly require foreign language instructors to act as proactive "ideological gatekeepers" to filter profound cultural biases. However, a teacher whose critical faculties and ideological sensitivity have been deskilled through habitual AI overreliance fundamentally lacks the meta-cognitive vigilance required to perform this cultural gatekeeping. As Qin explores in the theoretical context of human-machine symbiosis, educators must actively manage this technological interaction to maintain their authority.[7]

4.3 Theme 3: Global Technology vs. Local Values

The third structural tension identifies the profound conflict between the Western-centric epistemology inherent in global GenAI models and the deeply localized requirements of Chinese ideological and cultural education.

Cultural Representation Alienation and Epistemological Bias: Because the foundational architecture of major global GenAI models is overwhelmingly trained on WEIRD (Western, Educated, Industrialized, Rich, and Democratic) corpora, their latent

semantic networks inherently encode Western cultural norms. Consequently, when utilized to translate or generate discourse about unique Chinese cultural concepts, the resulting output frequently suffers from "cultural representation alienation."

From Passive Defense to Proactive Construction: Kong and Yang empirically investigate the application of Large Language Models, such as DeepSeek, in integrating Chinese culture into university-level foreign language teaching.[8] Rather than attempting to algorithmically filter out all biases, their research proposes a robust pedagogical model where educators serve as explicit "value guides." By intentionally deploying AI-generated cultural texts as objects of comparative study, teachers can foster advanced critical thinking.

5 Discussion and Limitations

While the synthesized CSSCI data highlights significant pedagogical shifts, a theoretical hypothesis emerging from the broader observation of the ecosystem is the role of institutional inertia. We hypothesize that deep, systemic adoption of GenAI into core educational functions may be heavily constrained by profound institutional inertia and enduring compliance-driven pressures of the academic environment.

Furthermore, this review is limited by its strict methodological focus on CSSCI-indexed journals from 2023 to 2025. While this ensures a high degree of pertinence to the local Chinese ideological context, it inevitably excludes valuable multi-modal and longitudinal data published in international high-impact journals. Future research should aim to conduct large-scale organizational studies to empirically test the hypothesized institutional barriers.

6 Conclusion and Implications: Bridging Academic Discourse and Classroom Reality

This comprehensive review maps the exploratory responses to GenAI within the Chinese EFL context. Empirical evidence indicates that current integration is largely gradual rather than structurally disruptive. Moving beyond academic hype toward a sustainable "Human-AI Symbiosis" relies on targeted interventions:

Structuring Human-in-the-Loop (HITL) Workflows: To mitigate the risk of cognitive offloading in mega-sized classes, studies emphasize collaborative intelligent feedback systems where teachers actively support students in interpreting and applying GenAI-generated feedback. **Cultivation of Critical AI Literacy (CAIL) and Open Box Pedagogy:** Responding to the threat of teacher deskilling involves a paradigm shift. Viewed through the international lens of Critical AI Literacy established by Pérez-Paredes, Curry, and Ordoñana-Guillamón [2], this review advocates for the integration of an "open box pedagogy" into Chinese FLT. By deliberately slowing down the learning process and demanding reflection, interpretation, and justification, this pedagogy creates "friction with purpose." It enables teachers to train students to critically interrogate the source and logic of AI outputs, acting as a crucial antidote to the "Gatekeeper-

Deskilling Paradox." Contextualization via Localized Knowledge Bases: To address cultural representation alienation, integrating Retrieval-Augmented Generation (RAG) architectures tied strictly to verified, localized Chinese cultural knowledge bases is essential.

References

1. Boumediene, Houda, Ouafa Ouarniki, and Pipit Rahayu. "ChatGPT Integration in Master Dissertation Writing: The Efficiency-Engagement Paradox." *Arab World English Journal*, 2024.
2. Pérez-Paredes, Pascual, Niall Curry, and Carlos Ordoñana-Guillamón. "Corpus Linguistics and AI in the Reconfiguration of Language Learning Ecologies." *Language Teaching*, 2026, pp. 1-15.
3. Xu, Linlin, Jinhuan Hu, and Yue Su. "EFL Students' Perceptions and Practices of AI-Assisted English Academic Writing." *Foreign Language World*, no. 3, 2024, pp. 51-58.
4. Wen, Qiufang, and Maocheng Liang. "Human-AI Interactive Negotiation Competence: ChatGPT and Foreign Language Education." *Foreign Language Teaching and Research*, no. 2, 2024, pp. 286-296.
5. Gao, Jie. "The Double-Edged Sword Effect of Generative Artificial Intelligence: Exploring Application Advantages and Potential Risks of DeepSeek in Foreign Language Education." *Journal of Contemporary Foreign Languages Studies*, vol. 25, no. 3, 2025, pp. 140-151.
6. Zhang, Ziyang, and Huaqing Hong. "ChatGPT and Foreign Language Teaching: Empowerment, Issues and Strategies." *Foreign Language World*, vol. 43, no. 2, 2023, pp. 38-44.
7. Qin, Yuan. "Exploration of Foreign Language Education Methods in the Context of Human-Machine Symbiosis: A Case Study on ChatGPT." *Technology Enhanced Foreign Language Education*, vol. 45, no. 2, 2023, pp. 24-29.
8. Kong, Lin, and Xiaorong Yang. "LLMs in Integrating Chinese Culture into Foreign Language Teaching." *Foreign Language World*, no. 1, 2025, pp. 29-37.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

