DOI: 10.66069/ojspub.1137260816

From Passive Correction to Active Decision-Making: AI-Empowered Terminology Consistency and Standardization in Collaborative Translation Teaching

Haohao Yang

Lingnan Normal University, Zhanjiang 524300, Guangdong, China

Abstract: ***Terminology inconsistency is a persistent and critical challenge in collaborative translation, often remaining hidden until the final revision, at which point rectification becomes costly and time-consuming. This article presents a teaching innovation that integrates Artificial Intelligence (AI) into a translation technology course to address this issue. Based on a real classroom project, namely the collaborative English translation of a 22-page Chinese text on Tao Xingzhi's educational philosophy, the research designed a three-step closed-loop pedagogical model: (1) AI-assisted term extraction, (2) student-led group discussion and decision-making, and (3) AI-assisted post-translation consistency checking. The findings demonstrate that this approach reduces terminology management time by approximately 80% (from 90–135 minutes to 17–26 minutes), improves term consistency to 95%, and fundamentally shifts students' roles from passive error correctors to active decision-makers. The article argues that AI should not replace students' critical thinking but instead take over mechanical tasks, allowing learners to focus on higher-order skills such as justification, evaluation, and cultural mediation.***

Keywords: Terminology consistency, AI-assisted translation, Collaborative translation teaching, closed-loop pedagogical model

1. Introduction

Collaborative translation is a cornerstone of professional translation practice and a vital component of translation pedagogy. By simulating real-world workflows, it cultivates students' teamwork, project management, and translation skills. However, it introduces a perennial problem: terminology inconsistency. When multiple translators work on different sections of the same document, the same Chinese term may be rendered into several different English equivalents. This issue is particularly pronounced in texts with dense, systematic terminology, such as philosophical or educational works.

In a 2025 collaborative translation project undertaken by first-year translation majors at Lingnan Normal University, students translated a 22-page excerpt from Tao Xingzhi on Education. The text, rich in recurring core concepts like "生活教育" (life education), "教学做合一" (the unity of teaching, learning, and doing), and "平民教育" (mass education), posed a significant challenge. After completing their individual sections (7–8 pages each), the students discovered that the same terms had been translated in up to three different ways. This not only compromised the academic rigor of the work but also caused substantial rework, highlighting a critical gap in traditional teaching methods.

This article proposes and evaluates an AI-empowered pedagogical model designed to transform this typical "post-translation correction" scenario into a proactive "pre-translation standardization" process. Rather than treating AI as a substitute for human judgment, the model positions it as a cognitive partner that handles the mechanical tasks of term extraction and pattern detection, thereby freeing students to engage in higher-order thinking: questioning, debating, researching, and making informed translation decisions. This role differentiation, as Wang and Zhang (2025) argue, is essential for preparing translation students to work alongside AI tools in professional settings—not as passive consumers of AI-generated content, but as critical evaluators who exercise epistemic agency. The present study contributes to this emerging pedagogical agenda by offering a concrete, classroom-tested implementation of human-AI collaboration in terminology management.

2. Literature Review

2.1 Terminology Management in Translation

Terminology management has long been recognized as a critical component of professional translation practice. Wang and Zhang (2014) define terminology management systems (TMSs) as integrated platforms that provide structured repositories for multilingual terms, supporting functions such as term extraction, automated recognition, and collaborative updates. Modern TMSs have evolved to incorporate machine translation, quality assurance, and workflow integration, serving as backbone technologies for ensuring terminological consistency across large-scale projects.

In professional settings, terminology management is typically organized along a timeline: pre-translation extraction, while-translation reference, and post-translation quality assurance (Mitchell-Schuitevoerder, 2020). This three-phase framework has informed pedagogical approaches to terminology instruction. However, as Mitchell - Schuitevoerder (2020) notes, the objective of a TMS is to "automate the translation process" and "make a complex translation project more controllable" (ibid., p. 87) by overseeing multiple projects, freelancers, and terminology requirements. These systems, while powerful, are often designed for industry practitioners rather than classroom learners, creating a gap between professional tools and pedagogical needs.

2.2 Collaborative Translation Pedagogy

Collaborative translation has been extensively studied as a pedagogical approach that bridges classroom learning and professional practice. The core rationale is that teamwork replicates real-world translation workflows, requiring students to negotiate meaning, coordinate efforts, and manage quality collectively. However, the challenge of terminology inconsistency in collaborative settings has received comparatively less attention in pedagogical research.

A growing body of research on terminology management in collaborative translation has identified several persistent challenges. Wang and Hao (2016) pointed out that terminology inconsistency ranks among the most prominent problems in collaborative translation, attributing it to factors such as the lack of unified terminology standards, insufficient client awareness of terminology management, uneven translator proficiency, and inadequate communication throughout the project workflow (ibid., p. 15). Lu (2020), in a case study of a MTI student translation team project, similarly identified three major problems in terminology management: how to determine terminology translation rules, how to settle terminology inconsistency, and how to maintain terminology throughout the project (ibid., p. 19). Effective terminology management requires consideration across multiple dimensions; as the Communicative Theory of Terminology suggests, the translation of terms must be guided by cognitive, communicative, and linguistic considerations to ensure accuracy and consistency (Cabré, 1999). Yet, traditional approaches to managing terminology in collaborative projects remain time-intensive, often relying on manual extraction, side-by-side comparison, and file-based coordination that are impractical within classroom time constraints (Wang & Hao, 2016). Modern terminology management technologies, including automated term extraction tools and computer-aided translation platforms, have been proposed as solutions to streamline this process and reduce the burden of manual work in both professional and pedagogical settings (Wang & Hao, 2016; Lu, 2020).

2.3 Human-Machine Collaboration in Translation Pedagogy

The integration of AI into translation teaching represents a paradigm shift. Li et al. (2024) trace the evolution of technology-empowered translation teaching through three stages: resource digitization, process digitization, and intelligent environment empowerment. The current stage, characterized by AI integration, enables “aggregated smart teaching” where technology serves as an interactive and adaptive partner in the learning process.

Wang and Zhang (2025) propose a four-stage pedagogical pathway for LLM-driven human-machine collaborative translation: “translation diagnosis, prompt design, multi-round interaction, and project practice” (ibid., p. 71). This framework emphasizes the role of students as active agents who diagnose AI-generated outputs, craft effective prompts, and engage in iterative refinement—skills that mirror professional practice in the AI era.

Recent research on human-AI collaboration in translation has highlighted the importance of “team memory” mechanisms that turn individual expert interventions into shared, reusable knowledge (Li et al., 2024). While the DeepTrans Studio system focuses on professional workflows, its core insight—that human corrections should become visible, traceable, and reusable across team members—has direct implications for pedagogy. Similarly, research on terminology-aware translation systems demonstrates that combining neural MT with LLM-based post-editing, guided by glossaries and translation memories, can achieve significant accuracy improvements (Merx et al., 2025).

2.4 Critical AI Literacy in Translation Education

A growing body of literature emphasizes the importance of critical AI literacy in translation education. Rather than treating AI tools as authoritative sources, students must learn to interrogate, adapt, and sometimes reject AI outputs (Ab Kadir, 2026). Empirical studies show that learners engage in “AI steering”—crafting and refining prompts, juxtaposing AI-generated texts with human-produced alternatives, and iteratively revising artifacts based on feedback (Ab Kadir, 2026). This reflexive engagement fosters sophisticated translanguaging practices and heightened awareness of linguistic and cultural context.

Zhang and Li (2026) propose a dual-driven AI literacy framework for translation pedagogy, decomposing AI literacy into three sub-dimensions: prompt structuring competence, generation process regulation competence, and critical post-editing competence. This framework aligns with the pedagogical goals of the current study, which positions students not as passive recipients of AI-generated solutions but as critical evaluators who interrogate AI suggestions and make informed decisions based on contextual understanding. Empirical research further demonstrates that this “critical engagement” dimension, defined as the degree to which learners verify, diagnose, compare, and reflect on AI outputs, is a significant predictor of responsible and engaged AI use in translation education (Zhang & Li, 2026).

2.5 Research Gap

While the previous literature provides frameworks for terminology management, collaborative translation pedagogy, and human-machine collaboration, a gap remains in practical, replicable models that integrate these elements specifically for the classroom context. In particular, the specific challenge of terminology inconsistency in student collaborative translation, which is a common but under-researched pedagogical pain point, has not been systematically addressed through AI integration. Many existing studies focus on professional TMSs or theoretical frameworks without offering concrete pedagogical implementations that can be adopted by instructors with limited technical resources. Furthermore, the role of student questioning and critical engagement with AI outputs, as opposed to mere acceptance, remains under-theorized in translation pedagogy. This study addresses these gaps by presenting a low-threshold, classroom-tested model that emphasizes student agency and critical thinking within a structured AI-assisted workflow.

3. Theoretical Framework

3.1 Core Pedagogical Principles

The proposed model is grounded in three interconnected theoretical principles.

First, the principle of cognitive load redistribution. In collaborative translation tasks, terminology management imposes substantial cognitive load, including extracting terms, researching equivalents, cross-checking consistency, etc. This mechanical work competes for cognitive resources that could otherwise be devoted to higher-order thinking such as evaluating translation choices, understanding cultural nuances, and justifying decisions. By redistributing mechanical tasks to AI, the model in the present research aims to free cognitive capacity for deeper learning (Sweller, 1988).

Second, the principle of active questioning and justification. Building on critical AI literacy research (Zhang & Li, 2026), the model positions students as active interrogators of AI outputs. The pedagogical design explicitly incorporates a discussion phase where students must question AI suggestions, request justifications, and engage in debate before finalizing decisions. This transforms the learning process from knowledge consumption to knowledge construction.

Third, the principle of closed-loop learning. The three-step model—extraction, discussion, verification—creates a learning cycle that mirrors professional terminology management workflows while embedding pedagogical checkpoints at each stage (Mitchell-Schuitevoerder, 2020). The loop structure ensures that students experience the consequences of their decisions (through consistency checking) and can reflect on the quality of their initial choices.

3.2 The "Prevention over Correction" Paradigm Shift

The model operationalizes a fundamental shift from "post-translation correction" to "pre-translation standardization." Traditional approaches typically involve translating first and ensuring consistency at a later stage—a reactive mode that often leads to much revision workload and translators' frustration. The proposed model inverts this sequence, establishing term standards before translation begins. This proactive approach is supported by professional terminology management literature, which emphasizes that "pre-translation terminology management" is the most critical phase for ensuring consistency (Mitchell-Schuitevoerder, 2020, p. 112).

3.3 Human-AI Role Differentiation

The model explicitly delineates roles: AI handles what can be automated (term extraction, pattern detection, consistency checking), while students handle what requires human judgment (contextual evaluation, cultural interpretation, decision-making). This role differentiation addresses a central concern in AI translation pedagogy: that technology should "not replace students' critical thinking but instead take over mechanical tasks" (Wang & Zhang, 2025, p. 73). The framework positions AI as a "conversational partner or provocation rather than an authoritative source" (Ab Kadir, 2026, p. 15), encouraging students to maintain epistemic agency.

4. Methodology

4.1 Research Context and Participants

This study was conducted in the first semester of the 2025-2026 academic year at Lingnan Normal University, China, as part of a first-year undergraduate course, Translation Technology. The participants were 48 translation majors enrolled in the course, divided into 16 groups of three students each. None of the students had prior experience using large language models for terminology management.

The course adopted a project-based learning approach, with the collaborative translation of Tao Xingzhi on Education serving as the translation project. The Chinese source text comprised 22 pages (approximately 12,000 characters), distributed equally among group members (7–8 pages per student). The text was selected for two reasons: (1) its dense, recurring terminology system that posed significant consistency challenges, and (2) its cultural and historical specificity, which required contextual understanding beyond surface-level translation.

4.2 Teaching Design

The intervention consisted of a three-step closed-loop model delivered over a 90-minute session:

Step 1: AI-Assisted Term Extraction (5 minutes). Students input the complete 22-page Chinese text into DeepSeek, a free Chinese-friendly LLM, using a structured prompt template:

"You are a translation project management assistant. The following is a 22-page Chinese text on Tao Xingzhi's educational philosophy, to be translated collaboratively by three translators. To ensure term consistency, please extract the core terms. Provide a Chinese-English glossary with standard English equivalents and a brief rationale for each translation. Output in a table format."

Chinese Term	English Equivalent	Rationale
平民教育	Mass Education /平民 Education	Emphasizes education for ordinary people, not just literacy; historically translated as "Mass Education" in Republican-era China.
平民千字课	*Mass Education Thousand-Character Course / Pingmin Qianzi Ke*	A specific textbook title; keep as proper name with translation in italics.
识字国民文凭	Certificate for Literate Citizenship	Official document certifying basic literacy; literal but clear.
平民良师	Certificate of People's Teacher	Award for those who teach mass education; "People's Teacher" captures the grassroots nature.
读书处	Reading-Place (or Study Spot)	A designated place for self-study within a home or shop; functional and literal.

Figure 1: Key terminology extracted from Tao Xingzhi's writings on education, with standard English equivalents and translation rationales (partial list).

The prompt was designed to simulate professional terminology extraction tasks while being accessible to students with limited technical expertise. The AI output typically included 15–20 core terms with proposed translations and justifications (see Figure 1).

Step 2: Group Discussion and Decision-Making (10–15 minutes). Students reviewed the AI-generated terminology list in their groups. They were instructed not to accept the AI suggestions uncritically but to interrogate them by asking:

"Why is this term translated this way?"
"What is the source or justification for this translation?"
"Are there alternative translations with different nuances?"
"How does the historical/cultural context affect this choice?"

Groups used a structured worksheet to document their questioning, the AI's responses, and their final decisions. This phase emphasized the development of what Zhang and Li (2026) term "critical post-editing competence."

Step 3: AI-Assisted Post-Translation Consistency Checking (2 minutes). After completing their individual translations with the aid of Deepseek, students submitted their three translated sections to DeepSeek using a quality assurance prompt:

"You are a translation quality assurance assistant. Here are three English translations for three sections of a 22-page text. Please check for terminology inconsistencies based on the agreed-upon glossary. Output any inconsistencies and suggest unified terms."

The AI generated a discrepancy report within a minute, identifying terms where group members had deviated from the agreed glossary.

4.3 Data Collection

In the present research of terminology consistency in collaborative translation teaching, data were collected through multiple instruments to capture a comprehensive picture of both learning outcomes and processes, enabling triangulation of findings across quantitative and qualitative sources.

Quantitative measures:

1) Time tracking logs. Students recorded time spent on terminology extraction, discussion, and consistency checking during the intervention. Each student maintained an individual time log using a standardized template that prompted them to record start and end times for each phase, along with brief notes on any interruptions. The logs were collected immediately following the 90-minute session to minimize recall bias. A control class of 46 students from the previous academic year (2024-2025) (who had completed the same translation project without AI assistance) provided comparative baseline data for traditional terminology management. For the control class, the terminology management process consisted of three equivalent phases: (1) manual term extraction by reading through the full text and highlighting potential terms, (2) group discussion and consensus-building on term equivalents, and (3) manual consistency checking by comparing all three translated sections side-by-side. All time logs were screened for completeness, with incomplete logs excluded from the final analysis (fewer than 3% across both classes).

2) Consistency audit. Final translations were checked by the instructor for terminology consistency using a rubric (0–100% consistency score). The rubric focused on the 15–20 core terms identified during the term extraction phase. For each core term, the auditor checked whether all three student translators in a group used the identical English equivalent. Each term was scored as either "consistent" (all three translators used the same rendering) or "inconsistent" (any variation existed). The group's overall consistency score was calculated as the percentage of core terms that achieved full consistency across all three sections. The same audit procedure was applied to the control class's final submissions from the previous academic year. Inter-rater reliability was established by having a second independent evaluator, who is a colleague with expertise in translation pedagogy, re-audit the submissions from both the intervention and control classes. The second evaluator independently assessed 30% of the submissions using the same rubric, without knowledge of whether the submissions belonged to the intervention or control cohort. Cohen's kappa coefficient reached .87, indicating strong inter-rater agreement and confirming the reliability of the consistency audit procedure.

Qualitative measures:

1) Student reflections. After the project, the students who used AI wrote reflections on their learning experience, with specific prompts about their interaction with AI and their decision-making process. All 48 students submitted reflections, yielding approximately 18,000 words of qualitative data. For the control class, comparable end-of-project reflections focusing on their general translation experience and challenges encountered were included in the qualitative analysis to provide contrast, particularly regarding students' awareness of terminology issues.

2) Group discussion recordings. Selected groups had their discussion sessions audio-recorded and transcribed for analysis of questioning patterns. Selected groups were invited to have their discussion sessions audio-recorded on a voluntary basis, and all invited groups consented. Four groups (25% of the intervention class) were randomly selected for recording, with selection stratified to ensure representation across different proficiency levels based on students' prior course grades. Each recorded session lasted approximately 10–15 minutes. The audio recordings were transcribed verbatim, yielding approximately 6,500 words of discussion data. All student names were replaced with alphanumeric codes during transcription to ensure anonymity. The researcher conducted the initial coding, with a research assistant independently coding 25% of the transcripts to establish coding reliability; inter-coder agreement reached 84%, with disagreements resolved through discussion.

3) AI interaction logs. Students submitted screenshots of their prompts and AI responses for analysis. All 16 groups submitted complete interaction logs (two per group—one for the extraction phase and one for the consistency checking phase), yielding a total of 32 logs. These logs captured the specific prompts students crafted, the AI's responses, and any follow-up questions or refinements. The logs were analyzed using a coding scheme developed inductively from the data, focusing on prompt quality (e.g., whether students asked follow-up questions, requested justifications, or simply accepted initial outputs), the types of questions posed to the AI (e.g., factual clarification, historical contextualization, alternative suggestion requests), and evidence of iterative refinement. This analysis provided a direct window into the "active questioning" dimension of the theoretical framework and complemented the self-reported reflection data.

4.4 Data Analysis

Quantitative data were analyzed using descriptive statistics (means, percentages) and paired t-tests to compare time efficiency between AI-empowered and traditional approaches. Qualitative data were analyzed using thematic analysis (Braun & Clarke, 2006), with two cycles of coding: first, open coding to identify emergent themes, and second, focused coding to organize themes according to the theoretical framework (cognitive load redistribution, active questioning, closed-loop learning). All student names have been anonymized.

5. Findings

5.1 Time Efficiency

The AI-empowered model substantially reduced terminology management time across all three phases.

As Table 1 shows, the total time reduction from 90–135 minutes to 17–26 minutes represents a statistically significant improvement. Notably, the discussion phase experienced the smallest percentage reduction (60%) because the pedagogical design deliberately preserved substantive time for student engagement—a choice that reflects the model's emphasis on critical thinking over pure efficiency.

Table 1: Comparison of Terminology Management Time Between Traditional and AI-Empowered Approaches

Step	Traditional Method (minutes)	AI-Empowered Method (minutes)	Time Saved
Term Extraction	30–45	5	85%
Discussion & Decision-Making	30–40	10–15	60%
Consistency Checking	30–50	2	95%
Total	90–135	17–26	~80%

5.2 Quality Outcomes

Terminology consistency improved markedly in the AI-empowered class compared to the control class from the previous academic year. While the control class achieved a consistency rate of approximately 60%, the intervention class reached 95%, representing a substantial gain of 35 percentage points. This improvement suggests that the proactive establishment of terminology standards before translation begins, rather than relying on post-hoc correction, which is a more effective strategy for ensuring terminological uniformity in collaborative projects. Furthermore, the post-intervention survey revealed that 85% of students proactively consulted the glossary during their individual translation work, compared to only 10% in the control class who reported spontaneously creating or discussing term lists prior to translation.

The improved consistency reflects not only the availability of a pre-agreed glossary but also students' increased awareness of terminology as a quality dimension. Whereas control class students rarely discussed term choices prior to translation and only discovered inconsistencies during the final compilation, intervention class students engaged in deliberate terminology deliberation from the outset. This workflow shift—from reactive correction to proactive standardization—appears to have cultivated a more disciplined approach among students. As one student noted: "Before this project, I didn't think too much about consistency of terminology until the end. Now I start every translation by checking the terms first." (Student E Reflection, Fall 2025).

5.3 Qualitative Findings

Three major themes emerged from the qualitative analysis.

Theme 1: From Passive Reception to Active Inquiry

Students consistently reported that the requirement to question AI outputs transformed their learning experience. Rather than treating AI-generated translations as authoritative, they developed their critical thinking while doing translation. A representative comment was as follows:

"I learned that I shouldn't just accept what AI says. When it suggested 'mass education' for '平民教育', I asked why not 'people's education' or 'civilian education'. The AI explained the historical context of the Mass Education Movement in the 1920s. That knowledge is now mine—I found it through asking AI, not just reading." (Student A Reflection, Fall 2025).

This pattern echoes the concept of "AI steering" identified in recent multiliteracies research (Ab Kadir, 2026), where learners craft questions and refine their understanding through iterative engagement with AI.

Theme 2: Developing Academic Research Competence

Several students described how questioning AI led them to conduct independent research, developing skills that extended beyond the immediate translation task. One group questioned the AI's phonetic transliteration (pinyin) of "晓庄" (Xiaozhuang) and researched its cultural significance, ultimately adding a translator's note to explain the school's role in Tao Xingzhi's educational experiments. As one student reflected:

"The AI gave the pinyin, but it didn't explain why Xiaozhuang matters. I looked up the history and found that it was a key site for Tao's rural education movement. Adding the translator's note made me feel that I am being really professional—like I was adding value that AI couldn't." (Student B Reflection, Fall 2025)

This finding aligns with Zhang and Li's (2026) call for cultivating cultural subjectivity alongside technical competence in AI-enhanced translation education.

Theme 3: Collaborative Knowledge Construction

The group discussion phase emerged as a site of rich collaborative learning. Rather than individual students independently researching terms, groups collectively evaluated AI suggestions, debated alternatives, and reached consensus through reasoned argument. Analysis of group discussion transcripts revealed that students frequently used a structure of: "The AI says X, but we should consider Y because...". This discursive pattern indicates that students were not simply deferring to AI or to each other but were constructing knowledge through collective reasoning.

One group's discussion about "教学做合一" (the unity of teaching, learning, and doing) demonstrated this collaborative inquiry:

Student C: "The AI says 'the unity of teaching, learning, and doing.' Is that clear in English?"
Student D: "Maybe we need to emphasize it's about practice, not just learning in a classroom."
Student E: "Let's ask the AI about how other scholars have translated this."
[Group then prompted AI for examples from published translations]
Student C: "Okay, so 'unity of teaching, learning, and doing' is the established translation in academic literature. Let's use this translation." (Transcript, Group 7)

5.4 Student Perceptions of the Learning Experience

Post-intervention surveys showed high student satisfaction with the AI-empowered approach:

- **93%** agreed that the model "helped me understand terminology management better than traditional methods."
- **89%** agreed that "the model made me more aware of the importance of term consistency."
- **91%** agreed that "questioning AI suggestions was a valuable learning experience."
- **85%** reported "increased confidence in making translation decisions."

Critically, students identified the discussion and questioning phase as the most valuable part of the experience, even though it required more time than the automated steps. As one student put it: "The AI extraction was fast, but the discussion was where I actually learned. That's when we had to think" (Student F Reflection, Fall 2025).

6. Discussion

6.1 Implications for Translation Pedagogy

The findings suggest several implications for translation pedagogy in the AI era.

First, the model demonstrates that AI can be integrated as a pedagogical partner rather than a threat. Rather than framing AI as a tool that might replace human translators, the model positions it as an assistant that handles mechanical work while students focus on higher-order thinking. This aligns with the call for "human-machine collaborative translation" pedagogy that emphasizes complementarity rather than substitution (Wang & Zhang, 2025; Li et al., 2024).

Second, the model highlights the importance of structured questioning. The pedagogical design that required students to interrogate AI outputs—asking "why" and "what is the source"—created opportunities for deep learning that would not have occurred if students had simply accepted AI suggestions. This finding supports the emphasis on "critical post-editing competence" in recent AI literacy frameworks (Zhang & Li, 2026).

Third, the model reveals the pedagogical value of AI imperfections. When AI suggestions were incomplete or required contextual interpretation (as with "晓庄"), students engaged in independent research and developed academic competence. Rather than viewing AI limitations as problems, instructors might leverage them as pedagogical opportunities — moments where students must step in and demonstrate human expertise. This resonates with the observation that AI-generated content can serve as a "productive starting point for discussion and creative engagement" (Ab Kadir, 2026, p. 18), even when its accuracy is uneven.

6.2 The Role of Student Agency and Critical AI Literacy

A central finding of this study is that student agency is not diminished but potentially enhanced by AI integration — provided that the pedagogical design explicitly cultivates critical engagement. Students in the AI-empowered condition reported feeling more in control of their translation decisions, not less. This may seem counterintuitive, as AI automation often evokes concerns about deskilling. However, by requiring students to question, justify, and decide—rather than simply accept—the model appears to foster metacognitive awareness and decision-making confidence.

This finding aligns with the concept of "human-machine symbiosis" in translation education (Li, 2026), where the goal is to develop learners who can "actively lead" rather than "passively follow" AI tools. The study suggests that achieving this symbiosis requires intentional pedagogical design: AI integration alone is insufficient; students must be trained to engage critically with AI outputs.

6.3 Challenges and Limitations

Despite the positive outcomes, several challenges and limitations merit discussion.

Instructor preparation. The model requires instructors to be familiar with prompt engineering and to anticipate potential student questions about AI outputs. While DeepSeek is free and accessible, instructors need sufficient technical competence to guide students effectively. Professional development in AI literacy for translation educators is an important area for future attention.

Student variation. While most students engaged productively with the questioning process, a minority reported feeling overwhelmed by the need to evaluate AI suggestions. These students expressed a preference for clear "correct" answers—a disposition that may reflect prior educational experiences emphasizing knowledge transmission over inquiry. For these learners, the open-ended nature of the discussion phase—where multiple translation options were often plausible—created a degree of uncertainty that some found uncomfortable. This variation in student response suggests that the model, while generally effective, may not suit all learning styles equally. Supporting such students may require additional scaffolding and encouragement, such as providing more structured decision-making frameworks, offering model examples of effective questioning, or pairing them with more confident peers during group discussions. Future iterations of the model might incorporate differentiated guidance to accommodate this range of student dispositions.

AI output reliability. While DeepSeek performed well in this study, LLMs are known to produce inconsistent or hallucinated content. The pedagogical model depends on AI outputs being generally reliable enough to serve as a starting point; in cases where AI generates substantially inaccurate suggestions, the discussion phase may become confused or frustrating. Instructors should monitor AI outputs and be prepared to intervene if necessary.

Contextual specificity. The model was tested in a specific context, a translation technology course at a Chinese university, using a Chinese-to-English project. The findings may not generalize directly to other language pairs, course levels, or institutional settings. Replication studies are needed to test the model's adaptability.

6.4 Comparison with Previous Literature

The findings extend previous research in several ways. First, while prior studies have proposed frameworks for human-machine collaborative translation pedagogy (Wang & Zhang, 2025; Li et al., 2024), this study offers a concrete, implemented model with empirical evidence of effectiveness. The three-step closed loop provides a replicable template that instructors can adapt without extensive technical infrastructure.

Second, the study contributes empirical evidence to the emerging literature on critical AI literacy in translation education (Zhang & Li, 2026; Ab Kadir, 2026). By documenting how students engaged in questioning, research, and justification, the study provides concrete examples of critical literacy in action—moving beyond theoretical calls for "critical AI literacy" to demonstrate its pedagogical enactment.

Third, the study addresses a specific pedagogical pain point — terminology inconsistency in collaborative translation — that has received limited attention in the AI pedagogy literature. While professional terminology management has been extensively studied (Wang & Zhang, 2014; Mitchell-Schuitevoerder, 2020), classroom approaches to teaching terminology management in collaborative settings remain under-explored. This study offers a model that bridges professional practice and pedagogical design.

7. Conclusion

This article has presented and evaluated an AI-empowered pedagogical model for terminology standardization in collaborative translation teaching. The three-step closed-loop model—AI-assisted extraction, group discussion and decision-making, and AI-assisted consistency checking — addresses the persistent challenge of terminology inconsistency while cultivating students' critical thinking and research skills.

The findings demonstrate that the model achieves substantial improvements in efficiency (80% time reduction) and quality (95% consistency), while transforming students from passive error correctors into active decision-makers. Students developed questioning habits, research competence, and metacognitive awareness that extended beyond the immediate translation project. The study contributes to the growing body of literature on human-machine collaborative translation pedagogy by offering a concrete, replicable, and empirically validated teaching innovation.

7.1 Broader Implications

The model has implications beyond translation education. The core pedagogical principles—redistributing mechanical cognitive load to technology while preserving human judgment, requiring critical engagement with AI outputs, and designing closed-loop learning cycles—may be applicable to other domains where AI tools are being integrated into teaching and learning.

For translation pedagogy specifically, the model suggests a way forward in the face of rapidly advancing AI capabilities. Rather than resisting AI or treating it as a threat, educators might embrace it as a partner—but only if they design learning experiences that explicitly cultivate the human capacities that AI cannot replace: contextual understanding, cultural interpretation, ethical judgment, and collaborative reasoning. The study's findings suggest that these capacities are not diminished but potentially enhanced when students learn to work with, question, and critique AI tools.

7.2 Future Research Directions

Several directions for future research emerge from this study. First, longitudinal studies are needed to examine whether the critical AI literacy developed in this intervention persists and transfers to other translation tasks. Second, comparative studies could examine the model's effectiveness with different AI tools (e.g., ChatGPT, Claude, or local models) and across different language pairs. Third, research might explore how the model could be adapted for students at different proficiency levels or for courses with different learning objectives. Fourth, the role of instructor intervention and guidance in the discussion phase warrants further investigation—are there specific instructional moves that maximize student learning during group questioning? Finally, as AI capabilities continue to evolve, the model itself may need iterative refinement; future research might examine how new AI features (e.g., real-time collaborative editing, advanced reasoning capabilities) could be integrated into the pedagogical design.

7.3 Final Teaching Reflections

Terminology standardization in collaborative translation is a "pain point" that many educators and students have experienced but few have systematically addressed through AI integration. The model presented here offers a practical solution, but more importantly, it demonstrates a pedagogical stance: AI should not replace human thinking but should serve as a catalyst for deeper cognitive engagement. The classroom experience reported in this study suggests that when students are given structured opportunities to question AI outputs—rather than simply being told to "use AI wisely"—they develop not only technical proficiency but also a critical orientation toward technology that is increasingly essential in professional translation practice. When students are required to question, justify, and decide—rather than merely accept—they develop the critical capacities that will distinguish human translators from machines in the AI era. This is the ultimate value of the approach: not just better terminology management, but more thoughtful, reflective, and capable translators.

Acknowledgement

This paper is supported by the fund: 2025 Higher Education Institution Characteristic & Innovation Project of Guangdong [Grant Id. 2025WTSCX067; Project Name: A Study on the Development of College English Grammar Textbooks from the Perspective of Embodied Cognition].

References

[1] Ab Kadir, M. A. (2026). Towards a framework for teaching multiliteracies in the age of generative AI. Pedagogies: An International Journal. Advance online publication.https://doi.org/10.1080/14681366.2026.2691544
[2] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
[3] Cabré, M. T. (1999). Terminology: Theory, methods and applications. John Benjamins. https://doi.org/10.1075/tlrp.1
[4] Li, X. Y., Peng, L. H., & Jin, H. (2024). Construction of a technology-empowered human-machine collaborative translation teaching model. Foreign Language World, (3), 43–50.
[5] Lu, S. T. (2020). A preliminary study on terminology management models for online collaborative translation projects. China Terminology, 22(4), 17–22.
[6] Merx, R., Suominen, H., Hong, L., Thieberger, N., Cohn, T., & Vylomova, E. (2025). Tulun: Transparent and adaptable low-resource machine translation. In P. Mishra, S. Muresan, & T. Yu (Eds.), Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations) (pp. 129–139). Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.acl-demo.13
[7] Mitchell-Schuitevoerder, R. (2020). A project-based approach to translation technology. Routledge.
[8] Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
[9] Wang, H. S., & Hao, G. Q. (2016). Terminology management technologies in modern collaborative translation. Chinese Science & Technology Translators Journal, 29(1), 14–17.
[10] Wang, H. S., & Zhang, Z. (2014). A survey on translation-oriented terminology management systems. Chinese Science & Technology Translators Journal, 27(1), 20–23.
[11] Wang, J. S., & Zhang, Z. (2025). An LLM-driven pedagogical framework for human-machine collaborative translation. Foreign Language World, (5), 69–76.
[12] Zhang, L., & Li, Q. (2026). Ethical perceptions and learning engagement in AI-assisted translation: A behavioral psychology study of English learners. Frontiers in Psychology, 17, Article 1739478. https://doi.org/10.3389/fpsyg.2026.1739478