Región Científica

e-ISSN: 2954-6168
Región Científica. 2026. 5(2), 2026613
doi: 10.58763/rc2026613

Review Article

From the subjunctive to the ergative: A narrative review of AI mediation in Spanish-Hindi literary translation and its impact on translation competence in university teaching

Del subjuntivo al ergativo: Una revisión narrativa de la mediación por IA en la traducción literaria español-hindi y su impacto en la competencia traductora en la enseñanza universitaria

Sabyasachi Mishra[1]

ABSTRACT

Introduction: The expansion of artificial intelligence has transformed translation and translator training, especially through neural networks and large-scale language models. In literary translation, these advances pose additional challenges due to the need to preserve linguistic, pragmatic, cultural, and stylistic dimensions. The aim was to critically analyze the mediation of AI in literary translation, with a special focus on Spanish–Hindi and its educational implications.

Methodology: A narrative, qualitative, interdisciplinary, and diachronic review was conducted for the period 1990–2026. The final corpus consisted of 27 studies selected through purposive and theoretical sampling, organized around five axes: linguistic divergence, pragmatic-cultural transfer, technological evolution, translator competence and cognitive load, and voice, authorship, and agency.

Results: The evidence showed that technological advances do not eliminate linguistic divergences nor guarantee stylistic preservation. LLMs expand creative possibilities, but they can also introduce unjustified expansions. Furthermore, technological development shifts human intervention from correction to evaluation, verification, and decision-making. In education, this process demands strengthening critical AI literacy, metacognitive regulation, and argumentative skills.

Conclusions: Spanish-Hindi literary translation remains insufficiently studied. Specialized corpora, empirical evaluations, and educational designs that conceive of AI as a heuristic mediator under human governance are needed.

Keywords: Artificial intelligence; Higher education; Linguistics; Literary analysis; Multilingualism.

JEL Classification: I20, I21, I23.

RESUMEN

Introducción: La expansión de la inteligencia artificial ha transformado la traducción y la formación de traductores, especialmente mediante sistemas neuronales y modelos de lenguaje de gran escala. En traducción literaria, estos avances plantean desafíos adicionales por la necesidad de preservar dimensiones lingüísticas, pragmáticas, culturales y estilísticas. El objetivo fue analizar críticamente la mediación de la IA en la traducción literaria, con especial atención al español–hindi y a sus implicaciones educativas.

Metodología: Se desarrolló una revisión narrativa, cualitativa, interdisciplinaria y diacrónica para el periodo 1990–2026. El corpus final estuvo conformado por 27 investigaciones seleccionadas mediante muestreo intencional y teórico, organizadas en cinco ejes: divergencia lingüística, transferencia pragmático-cultural, evolución tecnológica, competencia traductora y carga cognitiva, y voz, autoría y agencia.

Resultados: La evidencia mostró que los avances tecnológicos no eliminan las divergencias lingüísticas ni garantizan la preservación estilística. Los LLM amplían las posibilidades creativas, pero también pueden introducir expansiones no justificadas. Asimismo, el desarrollo tecnológico desplaza la intervención humana desde la corrección hacia la evaluación, verificación y toma de decisiones. En educación, este proceso exige fortalecer alfabetización crítica en IA, regulación metacognitiva y capacidad argumentativa.

Conclusiones: La traducción literaria español–hindi permanece insuficientemente estudiada. Se requieren corpus especializados, evaluaciones empíricas y diseños educativos que conciban la IA como mediador heurístico bajo gobernanza humana.

Palabras clave: Análisis literario; Enseñanza superior; Inteligencia artificial; Lingüística; Multilingüismo.

Clasificación JEL: I20, I21, I23.

Received: 01-06-2026 **Revised:** 22-07-2026 **Accepted:** 15-08-2026 **Published:** 31-08-2026

Editor: Alfredo Javier Pérez Gamboa

[1]Vellore Institute of Technology. Vellore, India.

Cite as: Mishra, S. (2026). From the subjunctive to the ergative: A narrative review of AI mediation in Spanish-Hindi literary translation and its impact on translation competence in university teaching. *Región Científica, 5*(2), 2026613. https://doi.org/10.58763/rc2026613

INTRODUCTION

Translation is undergoing a technological transformation that has simultaneously altered the available tools, professional workflows, and the competencies expected of those who practice or learn this activity. The progressive incorporation of statistical systems, neural machine translation, and, more recently, large-scale language models has shifted the use of artificial intelligence from relatively delimited linguistic transfer tasks toward interactive scenarios in which a system can translate, reformulate, explain choices, respond to instructions, and produce alternatives. This evolution has reinforced the interest in understanding not only how much can be automated, but also under what linguistic, textual, and educational conditions it is appropriate to do so. In this context, He et al. (2022) identified a sustained expansion of research on teaching translation technologies, with particular international interest in machine translation and post-editing, while He and Tao (2022) explicitly develop the notion of translation technological thinking competence and integrate technological awareness, learning, application, evaluation, creation, and problem-solving.

The emergence of LLMs has increased the complexity of this discussion, since, unlike systems specifically designed for translation, generative models acquire translation capabilities through general-purpose multilingual training and can modulate their behavior through instructions. Peng et al. (2023) demonstrated that variables such as domain information, task formulation, and generation parameters modify ChatGPT's translation performance, introducing an interactive dimension that was not present with the same intensity in previous paradigms.

Similarly, Wang et al. (2023) showed that LLMs can leverage documentary context to address discursive phenomena beyond the isolated sentence. However, this flexibility does not imply uniform performance, as observed in the study by Manakhimova et al. (2023), who found significant differences depending on the linguistic directions and phenomena evaluated, with persistent difficulties in aspects such as idioms, semantic roles, and certain voice configurations.

The unequal availability of resources constitutes another structural condition of the problem. In an evaluation of 204 languages, Robinson et al. (2023) observed that the level of linguistic resources was one of the main factors associated with the relative performance of ChatGPT, particularly unfavorable for numerous languages with fewer resources. This asymmetry is relevant because general progress in multilingual models can mask important differences between specific language combinations.

Recent research with Indic languages shows precisely an expansion of models, assessment suites, and specialized strategies, but also the need to continue developing specific resources. Rajpoot et al. (2024) addressed multimodal translation into Hindi and other Indic languages with fewer resources using linguistic models and knowledge distillation. For their part, Zerva et al. (2024) incorporated English–Hindi into multidomain quality assessment suites and automatic post-editing. These efforts reflect a growing computational interest in Hindi, although much of the experimental infrastructure continues to revolve around English as a source or target language.

The problem becomes particularly complex when the technology is applied to literary translation. Unlike domains where adequacy can be assessed primarily through terminological accuracy and information transfer, literature involves discourse relations, ambiguity, characterization, register variation, metaphor, cultural references, and stylistic patterns that depend on units larger than the sentence. For example, Karpinska and Iyyer (2023) found that providing a Language Learning Manager (LLM) with complete literary fragments could improve translation compared to sentence-by-sentence processing in different language pairs, by reducing certain errors and promoting stylistic consistency. However, they also documented critical errors and omissions that justify maintaining human intervention. This evidence is especially significant because it positions context as a necessary, but not sufficient, condition for preserving the complexity of the literary work.

In a combination like Spanish-Hindi, two linguistic traditions with substantial differences in their morphosyntactic, semantic, and pragmatic mechanisms converge. The challenge lies not only in replacing lexical units but also in reconstructing relationships in the target language that can be distributed differently across morphology, syntax, and discourse context. At the same time, contemporary scientific literature on AI and translation focuses predominantly on languages with high availability or on combinations centered around English. Therefore, it is pertinent to approach Spanish-Hindi not as a mere peripheral application of advances made for other languages but as a research space capable of testing the transfer of systems developed under very different data availability conditions. Recent global literature demonstrates that the performance of language-based translation (LBT) systems depends as much on the linguistic direction as on the phenomena being evaluated, so peer inferences must be made with caution

(Manakhimova et al., 2023; Robinson et al., 2023).

Technological transformation also directly affects higher education, as the issue is no longer limited to incorporating digital tools into traditional subjects, but rather requires reconsidering which skills remain fundamental when a growing part of the process can be mediated by automated systems. In this context, Ehrensberger-Dow et al. (2023) propose expanding machine translation literacy to include an understanding of its foundations, benefits, limitations, and risks, so that translators and trainers can decide when and how to use it. Along the same lines, Prieto Ramos (2024) argues that models of translation competence need to be updated in light of the impact of AI on professional methods and validates this need based on information from professionals in international organizations. These approaches allow us to consider technological literacy as an integrated dimension of translation competence and not as a peripheral skill added later.

Curriculum updates also present an institutional and professional dimension, as Al-Batineh and Al Tenaijy (2024) demonstrated through their analysis of job postings and training programs. Their research revealed that technological demands in the market can evolve more rapidly than university curricula, generating mismatches in areas such as computer-assisted translation, localization, and machine translation. In Spain, González-Pastor (2024) also examined the relationships between translator training, machine translation, and professional needs, highlighting the necessity of maintaining a continuous connection between academia and technological evolution. From a broader perspective, Tian (2024) argues that digital intelligence opens up possibilities for personalized training and more flexible feedback, but also necessitates addressing issues related to data quality, technological dependence, ethics, and teacher-student interaction.

The first studies specifically focused on GenAI within translator training confirm the need to analyze these technologies as educational mediators and not simply as productivity tools. Specifically, Kwok et al. (2025) studied post-editing with GPT in student translations and observed modifications in the lexical and syntactic complexity of the texts, although not uniformly, leading them to emphasize the importance of critical AI literacy linked to linguistic and instrumental competence. This type of study broadens the research agenda because, in addition to asking whether AI improves a textual product, it demonstrates the need to examine which cognitive operations, decisions, and learning it promotes or displaces during training.

Within this framework, Spanish-Hindi literary translation offers a particularly fertile case for integrating three fields that frequently advance in parallel: translation studies, computational linguistics, and university pedagogy. The need for this convergence stems from the fact that a system can exhibit high overall performance and, at the same time, prove insufficient for a particular linguistic phenomenon. Likewise, a system can generate fluent formulation that requires specialized cultural or stylistic evaluation and can be technologically useful without its incorporation into the classroom automatically producing higher-quality learning. Consequently, a review focused exclusively on automated metrics or comparisons between systems would be insufficient to understand the implications of the problem.

Based on these considerations, the purpose of this article is to critically analyze the mediation of artificial intelligence in literary translation, with special attention to Spanish–Hindi, and to examine its implications for university-level translator training. Through an interdisciplinary narrative review with a diachronic perspective, five key themes are articulated: the morphosyntactic and semantic divergences associated with Hindi; pragmatic and cultural transfer; the evolution from rule-based systems and statistical models to NMT and LLM; the transformations of post-editing, cognitive load, and translator competence in educational contexts; and issues related to voice, authorship, creativity, and human agency. The aim is not to establish a benchmark for Spanish–Hindi performance, but rather to critically organize the available evidence to define the state of knowledge, identify the problems that require specific research, and provide a relevant conceptual framework for teaching literary translation in increasingly AI-mediated settings.

METHODOLOGY

Review design

This study was conducted as a qualitative narrative review aimed at critically examining the role of machine translation and artificial intelligence tools in literary translation between Spanish and Hindi, as well as their implications for university-level translator training. The review's purpose was not to establish aggregate performance comparisons between systems or to rank technologies based on automated metrics, but rather to reconstruct the

main linguistic, technological, stylistic, pedagogical, and epistemological transformations that accompany the progressive incorporation of artificial intelligence into translation practice.

The review adopted a diachronic, interdisciplinary, and critical perspective. The diachronic dimension allowed for tracing the evolution from early machine translation systems based on rules and interlinguistic architectures to statistical models, neural machine translation, multilingual systems, and current extensive language models. The interdisciplinary dimension brought together contributions from translation studies, contrastive linguistics, computational linguistics, natural language processing, literary translation, and translation pedagogy. Finally, the critical dimension allowed for the analysis not only of the tools' technical performance but also of their effects on translator agency, stylistic preservation, cognitive load, translation competence, and authorship.

The search period was established between 1990 and 2026, with the aim of covering both the history of machine translation applied to typologically divergent languages and recent developments associated with neural models and generative systems. Although the search period began in 1990, the oldest document ultimately included in the analytical corpus dates from 1995.

Document search strategy

The search was conducted using a phased strategy that combined bibliographic databases, specialized repositories, and scientific sources specific to computational linguistics. The primary databases considered were Scopus, Web of Science, Google Scholar, ACL Anthology, MLA International Bibliography, and ERIC, supplemented by targeted searches in journals specializing in translation, linguistics, artificial intelligence, and higher education.

Due to the relevance of peer-reviewed scientific communications in computational linguistics, papers published in the proceedings of specialized conferences, particularly ACL Anthology, EAMT, Machine Translation Summit, and WMT, were also considered, provided they presented sufficient methodological information and contributed evidence directly related to the research problems. The search equations were constructed by combining four core concepts: language pair, translation technology, literary translation, and translator training. Among the main combinations, expressions equivalent to:

Spanish-Hindi machine translation, Spanish Hindi translation, Spanish to Hindi translation, Hindi machine translation, Hindi neural machine translation, Hindi large language models translation, literary machine translation, AI-assisted literary translation, literary translation LLM, post-editing literary translation, Spanish literary machine translation, translator education artificial intelligence, translation competence AI, student translators GenAI, cognitive load post-editing, translator voice machine translation, Hindi ergativity machine translation, Hindi aspect translation, politeness translation, cultural markers machine translation y creative phraseology translation.

The search was complemented with specific terms linked to the linguistic and stylistic phenomena of interest, such as *ergativity*, *subjunctive*, *aspect*, *discourse markers*, *politeness*, *āp/tum/tū*, *over-translation*, *under-translation*, *metaphor*, *voice*, *style*, *creativity*, *literary post-editing* y *cultural pragmatics*.

Sampling and formation of the analytical corpus

A purposive sampling method based on theoretical relevance was used, aimed at selecting documents best suited to answer the review questions. The sample was not defined by criteria of statistical representativeness, but rather by conceptual density, disciplinary diversity, proximity to the object of study, and interpretive sufficiency.

The search revealed a limited availability of research specifically focused on Spanish-Hindi literary translation mediated by artificial intelligence. For this reason, the corpus was not artificially restricted to studies exclusively related to this language pair. Instead, a hierarchical structure of evidence was constructed that allowed Spanish-Hindi to remain the epistemological center of the study while simultaneously integrating research from related fields when it offered transferable elements for understanding the phenomena analyzed.

The final analytical corpus consisted of 27 publications. The increase in the sample size from the initial estimate of approximately 20 documents was due to the thematic fragmentation of the literature and the need to ensure sufficient coverage of the linguistic, technological, literary, pedagogical, and ethical dimensions of the problem. Expanding the sample to a maximum of 30 works was not deemed necessary, as the 27 selected documents provided sufficient conceptual coverage and began to show recurrence in the central interpretive categories. The selection

was organized according to five levels of evidence (table 1).

Table 1.
Levels of organization and evidence

Level	Nature of the evidence
I. Direct Evidence	Studies specifically related to Spanish-Hindi translation or multilingual systems that integrate both languages.
II. Proximate Linguistic Evidence	Research on machine translation from Hindi and relevant structural phenomena, such as ergativity, aspect, syntactic divergence, pragmatic marking, politeness, and constituent order.
III. Comparative Literary Evidence	Studies on literary translation using neural machine translation, generative systems, or extensive language models in distinct language pairs with comparable stylistic, cultural, or pragmatic problems.
IV. Pedagogical and Cognitive Evidence	Research on post-editing, cognitive effort, translator training, translation competence, and the incorporation of artificial intelligence in educational contexts.
V. Ethical and Epistemological Evidence	Studies on voice, authorship, creativity, agency, idiolect, professional autonomy, and the human-machine relationship in literary translation.

Proximity to the Spanish-Hindi pair was used as a hierarchical selection criterion, but not as a rigid numerical quota. When direct evidence proved insufficient, bridging studies were used whose value was supported by clearly identifiable linguistic, literary, technological, or pedagogical correspondences.

The 27 selected documents were analytically distributed across three main domains: 11 studies primarily related to Hindi and the evolution of machine translation systems; 13 studies linked to literary translation, style, creativity, Spanish, and artificial intelligence; and 3 studies specifically focused on training, cognition, post-editing, and translation competence. This distribution was used only as an initial organizational criterion, as each publication could simultaneously contribute to different dimensions of analysis (Tables 2, 3, and 4).

Tabla 2.
Hindi Core and the Evolution of Machine Translation — 11 Studies

N°	Selected study	Type/object	Specific contribution	Questions
1	Sinha et al. (1995), *ANGLABHARTI: A Multilingual Machine Aided Translation Project on Translation from English to Indian Languages*	Rule-based system/ pseudointerlingua	This is the historical starting point. It allows us to reconstruct how attempts were made early on to resolve the SVO→SOV divergence and the generation towards Hindi using rules and human post-editing. DOI 10.1109/ ICSMC.1995.538002.	Q1, Q3
2	Dave, Parikh & Bhattacharyya (2001), *Interlingua-based English-Hindi Machine Translation and Language Divergence*	EnglishHindi; Machine Translation 16(4)	Fundamental to the notion of linguistic divergence. It analyzes syntactic, lexico-semantic, and generational divergences between structurally different languages.	Q1, Q2
3	Sinha & Jain (2003), *AnglaHindi: an English to Hindi machine-aided translation system*	Hybrid system rules + examples + statistics	It documents the transition from a purely regulated paradigm to an initial hybridization, while maintaining human intervention.	Q1, Q3
4	Sinha, Mahesh & Thakur (2005), *Translation divergence in English-Hindi MT*	EnglishHindi divergence	Very useful for systematizing those correspondences that do not allow direct transfer between linguistic systems.	Q1
5	Sinha & Thakur (2005), *Handling ki in Hindi for Hindi-English MT*	particle/complementizer *Ki*	Excellent micro-study to demonstrate that the same Hindi unit can fulfill different syntactic-semantic functions and require contextual disambiguation.	Q1, Q2
6	Bojar, Straňák & Zeman (2010), *Data Issues in English-to-Hindi Machine Translation*	SMT, English→Hindi corpus	This introduces a critical axis for our article: having a lot of data is not enough; domain, noise, alignment, and overlaps radically affect quality.	Q1, Q3

7	Goyal, Mishra & Sharma (2020), *Linguistically Informed Hindi-English Neural Machine Translation*	Transformer/NMT	It shows that the complex morphology and relatively free ordering of Hindi continue to be problems even in NMT and that POS, lemma, and morphological features improve the system.	Q1, Q3
8	Fan et al. (2021), *Beyond English-Centric Multilingual Machine Translation*	M2M-100	A turning point towards many-to-many translation without a mandatory pivot in English. This is crucial for explaining technologically how a SpanishHindi direction becomes viable.	Q1, Q3
9	Goyal et al. (2022), *The FLORES-101 Evaluation Benchmark for Low-Resource and Multilingual Machine Translation*	Benchmark multilingüe	It includes Spanish and Hindi within a professionally translated benchmark and allows for many-to-many evaluation; it highlights the difference between having a general evaluation and having a specialized literary corpus.	Q1, Q3
10	NLLB Team (2024), *Scaling neural machine translation to 200 languages*	NLLB-200	It represents the maturation of the massively multilingual NMT and the reduction of the bias towards high-resource languages, but continues to depend heavily on the availability and quality of bitexts.	Q1, Q3
11	Bhattacharjee, Gain & Ekbal (2024), *Domain Dynamics: Evaluating Large Language Models in English-Hindi Translation*	LLM, English→Hindi	Especially valuable because it tests LLM across various domains, including literary/religious material, showing that performance depends on the domain and that generalist fluency does not guarantee specific sensitivity.	Q1, Q2, Q3

Table 3.
Literary translation, Spanish, style and creativity — 13 studies

N°	Selected study	Type/object	Specific contribution	Questions
12	Toral & Way (2015), *Machine-assisted translation of literary text: A case study*	Assisted literary translation	This foundational work challenges the supposed incompatibility between literature and MT; it emphasizes the need to preserve the reading experience, not just the propositional content.	Q2, Q3, Q5
13	Toral & Way (2018), *What Level of Quality Can Neural Machine Translation Attain on Literary Text?*	Novels, NMT vs. PBSMT	It trains systems with over 100 million literary words and demonstrates improvements in NMT compared to SMT, but also a substantial gap compared to professional translation.	Q2, Q3
14	Toral, Wieling & Way (2018), *Post-editing Effort of a Novel With Statistical and Neural Machine Translation*	Experiment with literary translators	Six professional translators, English→Catalan novel. It provides temporal, technical, and cognitive measures of post-editing. It is central to connecting literature and human effort.	Q3, Q4
15	Kenny & Winters (2020), *Machine translation, ethics and the literary translator's voice*	NMT, translator's voice	One of the core points of the article: the textual voice of the professional translator appears attenuated when working through post-editing versus translation from scratch.	Q3, Q5
16	Macken, Vanroy, Desmet & Tezcan (2022), *Literary translation as a three-stage process: machine translation, post-editing and revision*	English→Dutch	It allows analyzing literary translation as a flow MT→possession→revision and distinguishing what modifications each stage introduces.	Q2, Q3, Q4
17	Ibáñez Moreno & Domínguez Mora (2025), *Google Translate versus DeepL in Spanish to English translation of Don Quixote*	Spanish→English; Don Quijote	An exceptionally relevant source for the Spanish component: it evaluates NMT in a canonical text and phraseological/collocational phenomena of high literary complexity.	Q1, Q2, Q3

18	Du et al. (2025), *Optimising ChatGPT for creativity in literary translation: ... English into Dutch, Chinese, Catalan and Spanish*	LLM, includes Spanish	It analyzes granularity, temperature, and prompting in creative translation. ChatGPT can outperform DeepL in some settings, although it still falls short of human translation.	Q2, Q3, Q5
19	Noriega-Santiáñez & Corpas Pastor (2025a), *Measuring Creative Phraseology in Literature: Machine Translation Systems Versus Large Language Models*	English→Spanish, literary phraseology	Compare DeepL/Google with ChatGPT/Gemini on creative phraseology; allows working on the resilience of non-compositional and stylized units against automation.	Q1, Q2, Q3
20	Noriega-Santiáñez & Corpas Pastor (2025b), *Technology and GenAI adoption among literary translators in Spain*	Survey of Spanish professionals	It introduces the perspective of the actual translator: uses, resistance, and attitudes toward GenAI. Discomfort and distrust toward GenAI are much greater than toward general technological tools.	Q3, Q5
21	Tewari & Baghel (2026), *Stylistically-Aware Hindi-English Poetic Translation with mBART and LLM-Based Post-Editing*	Hindi→English Poetry, mBART50 + LLM	It is probably the work closest to the literary heart of the Hindi component: curated poetic corpus, stylistic labeling and post-editing with LLM for metaphor, emotion, tone and rhythm.	Q1, Q2, Q3, Q5
22	Huang & Cheung (2026), *Exploring AI's performance in literary autobiography translation: how closely do AI models match human translation*	NMT, LLM and human translation	It compares AI models of different natures with human translation in a literary genre and focuses precisely on depth and nuance.	Q2, Q3, Q5
23	Wang et al. (2026), *Workflow matters: Comparing human translators and multi-agent LLMs in literary translation*	Target; LLM agents	This is highly relevant because it shifts the question from «which model is better» to how the human-machine flow is organized, a central concept for our pedagogical interpretation.	Q3, Q4, Q5
24	Resende & Hadley (2026), *Extending Creativity: Large Language Models and the Practice of Poetry Translation*	LLM and poetry	It proposes prompting strategies in pre-translation and translation phases and conceptualizes the LLM as a possible extension of creativity, not necessarily as a replacement.	Q2, Q3, Q5

Table 4.
Training, cognitive load, and ethics — 3 studies

N°	Selected study	Type/object	Specific contribution	Questions
25	Taivalkoski-Shilov (2019), *Ethical issues regarding machine(-assisted) translation of literary texts*	Critical essay on translation studies	It establishes a link between quality, ethics, voice, multivocality, and textual propriety. This will be one of the foundations for defining the humanist approach to revision.	Q5
26	Rojo López, Vicente López & Hvelplund (2024), *Measuring cognitive effort in post-editing: an eye-tracking study comparing professional and student translators*	25 professionals + 27 students; English→Spanish	It directly compares professionals and students using eye tracking. This provides much stronger evidence than simple perception surveys for assessing cognitive load in training.	Q3, Q4
27	Mau, Wu & Feng (2026), *Mind the cognitive load gap: Student translators' cognitive demands and translation behaviors in GenAI-assisted legal translation*	35 students; GenAI; training	Although the domain is legal, it is methodologically transferable: it demonstrates different cognitive profiles, the risk of overconfidence, and the need to transform GenAI verification into metacognitive competence.	Q3, Q4

Inclusion and exclusion criteria

Academic publications were included if they met one or more of the following criteria: a) analyzing machine translation, neural translation, extended language models, or artificial intelligence applied to translation; b)

studying Hindi, Spanish, or linguistic phenomena relevant to understanding their differences; c) addressing literary translation, creativity, metaphor, phraseology, voice, register, pragmatics, or cultural transfer; d) examining post-editing processes and human-machine collaboration; e) studying translation competence, cognitive load, or translator training; and f) contributing relevant conceptual or empirical elements to the discussion of authorship, agency, and the professionalization of translation.

Scientific articles, papers published in peer-reviewed proceedings of specialized conferences, and other academic documents with sufficient methodological traceability were accepted. Duplicate publications were excluded, as were exclusively technical studies without linguistic, translation studies, or pedagogical implications; works focused on highly specialized domains without a reasonable possibility of transfer to the object of study; documents lacking sufficient methodological information; and publications whose connection to the research questions was tangential, and studies whose complete information did not allow for adequate analytical reading.

Automatic translation metrics, including BLEU, METEOR, TER, and COMET, were not used as inclusion criteria or as an independent mechanism for determining translation quality. When they appeared in the selected studies, they were interpreted as a complementary dimension and always in relation to semantic, pragmatic, stylistic, cultural, and human aspects.

Analytical corpus and sources of foundation

Methodologically, a distinction was made between the analytical corpus and sources of theoretical grounding. The 27 documents that comprised the sample constituted the units actually subjected to coding, comparison, and interpretation. In contrast, established models of translation competence, descriptive studies on Spanish and Hindi grammar, methodological references on narrative revision, and other conceptual works were used to support the interpretation of the results, but were not counted as units in the sample.

This distinction prevented the artificial inclusion in the corpus of documents used exclusively to define concepts or support methodological decisions.

Extraction and organization of information

Each of the 27 selected studies was read in its entirety and recorded in a qualitative analysis matrix designed to identify its main characteristics in a standardized manner. The matrix included the following categories: year and context of publication; source language and target language; translation technology used; text genre; nature of the corpus used; morphosyntactic phenomenon studied; semantic issues; treatment of cultural markers; treatment of politeness and register; presence of metaphor, phraseology, or stylistic devices; evaluation system used; type of human intervention; nature of post-editing; cognitive effort or load; implications for translation competence; educational applications; conception of human agency; treatment of voice or idiolect; main findings; and limitations acknowledged by the authors. Additionally, each study was coded according to its contribution to the five guiding questions of the review:

1. 1. Morphosyntactic and semantic gap.
2. Pragmatic and cultural transfer.
3. Technological evolution and post-editing.
4. Translation competence, teaching, and cognitive load.
5. Ethics, authorship, agency, and idiolect.

The cross-cutting category of epistemic stance toward artificial intelligence was also incorporated, intended to identify how each publication conceptualized the relationship between technology and translators. For this purpose, an interpretive continuum was used, encompassing predominantly instrumentalist positions (AI as an extension or support of human activity), collaborative perspectives (AI as a component of human-machine interaction), and critical or humanist positions (AI as a potential threat to the translator's autonomy, voice, creativity, or professionalization). This classification was not treated as a rigid dichotomy and allowed for the recording of intermediate positions.

Analysis strategy

The information was examined using a qualitative, abductive thematic analysis. The deductive component consisted of the five questions that structured the review, while the inductive component allowed for the identification of emerging categories derived from the comparative reading of the documents.

The analysis focused particularly on phenomena that could constitute critical points in Spanish-Hindi translation: verbal modality and subjunctive, tense and aspect, ergativity, diathesis, causality, constituent order, pronominal reference, proclisis and enclisis, forms of address, politeness, register, discourse markers, phraseology, metaphor, irony, polyphony, and culturally situated references.

The technological component examined the evolution of rule-based systems toward statistical, neural, multilingual, and generative architectures, paying special attention to the changes introduced in the distribution of tasks between machine and translator. In addition, post-editing was analyzed not only as a corrective operation, but also as a cognitive and stylistic practice capable of modifying how the translator interprets, evaluates, and reconstructs the text.

The pedagogical component analyzed the implications of this transformation for translation competence, especially regarding the ability to detect errors, evaluate automatically generated options, verify cultural appropriateness, justify stylistic decisions, regulate technological dependence, and develop metacognitive monitoring strategies.

Finally, the ethical and epistemological component examined the extent to which algorithmic mediation modifies the translator's authorship, agency, and voice, differentiating between situations in which artificial intelligence functions as a heuristic tool and those in which its use can lead to stylistic homogenization, idiolect reduction, or displacement of professional judgment.

The results were integrated through a critical narrative synthesis, seeking to establish relationships between technological changes, linguistic problems, and pedagogical transformations. Thus, the review was not limited to describing the evolution of the tools and sought to determine which aspects of literary translation continue to show resistance to automation, which decisions still require particularly sophisticated human intervention, and what skills the university translator must develop to interact critically with artificial intelligence systems.

RESULTS

From morphosyntactic divergence to the loss of semantic-aesthetic density

The first cross-cutting pattern identified shows that machine translation difficulties related to Hindi cannot be reduced to superficial differences in word order. Sinha et al. (2005) characterized the Hindi–English relationship as a space of simultaneously grammatical and extragrammatical divergences, with differences affecting the realization of argument structure, determination systems, morphology, verb constructions, particles, and certain sociocultural components. Among the most significant cases are constructions in which an argument realized as nominative in English is marked by a dative in Hindi, as well as differences in the expression of modality and aspect that do not admit uniform structural correspondences.

The difficulty increases when morphosyntactic distinctions are directly involved in the construction of meaning. Sinha et al. (2005) observed, for example, that certain Hindi passive constructions express values of non-volition whose recovery in English requires semantic reformulation, while particles, expressive words, and echo forms combine syntactic, semantic, and sociocultural functions that do not always have direct equivalents. The same authors showed that honorifics can be expressed through plural pronouns and verbal agreement, thus integrating grammatical information and social relations within a single linguistic choice. In a specific analysis of the particle ki, Sinha and Thakur (2005) identified functions such as complementizer, coordinating conjunction, marker of purpose, contrastive negation, and expression linked to clauses of possibility, whose disambiguation depends on both the syntactic structure and the semantic context.

The transition to neural machine translation did not render these particularities irrelevant. Goyal et al. (2020) noted that morphological complexity, relatively free ordering, and insufficient parallel data continue to limit the performance of Hindi-English systems. By incorporating lemmas, grammatical category labels, and morphological features into a Transformer model, the authors found improvements in both word-based and subword-based models; specifically, the configuration integrating linguistic information outperformed the word-based baseline by more than two BLEU points. These results suggest that neural representations can learn complex linguistic regularities, but do not necessarily render explicit knowledge of language structure redundant.

The development of massively multilingual models has considerably expanded the possibilities for direct translation between languages other than English. Fan et al. (2021) developed M2M-100 for 100 languages and

9,900 translation directions, achieving average improvements of more than ten BLEU points compared to an English-centric multilingual model in the non-English directions evaluated. Subsequently, the NLLB Team (2024) extended this approach to 200 languages and approximately 40,000 directions, with an average improvement of 44% compared to previous systems. However, the study itself attributed much of the performance disparity to the availability, diversity, and quality of parallel data, which is particularly problematic for languages with fewer resources. Consequently, the incorporation of Spanish and Hindi into contemporary multilingual architectures demonstrates technological feasibility, but does not constitute sufficient evidence of suitability for literary translation between the two languages.

The specific literary evidence available for Hindi reinforces this distinction between linguistic correctness and aesthetic preservation. Tewari and Baghel (2026) found that general systems could produce semantically plausible translations of Hindi–English poetry while simultaneously weakening metaphor, rhythm, tone, and emotional resonance. Adapting mBART50 using an annotated poetic corpus increased BLEU from 12.4 to 18.3 and BERTScore from 0.843 to 0.886; furthermore, post-editing using an LLM increased human-rated poetic fidelity from 3.7 to 4.5 and fluency from 3.9 to 4.7 on a five-point scale (Tewari & Baghel, 2026). This result indicates that reducing semantic error does not, in itself, guarantee the preservation of semantic-aesthetic density, as this requires modeling stylistic dimensions that go beyond propositional equivalence.

In contrast to this accumulation of evidence on Hindi-English and multilingual systems, none of the 27 studies analyzed experimentally examined the Spanish-Hindi literary translation of phenomena such as the subjunctive, tense-aspect interaction, ergativity, or the resolution of Spanish clitics. Therefore, the findings establish a consistent linguistic basis for considering its transfer problematic, but do not demonstrate how current systems resolve it in this pair. This lack of direct evidence constitutes, in itself, one of the main gaps identified by the review.

Pragmatic-cultural transfer: between neutralization, loss and generative expansion

The second pattern identified shows that the difficulties of AI-mediated translation are not limited to the transfer of propositional content, but extend to elements whose meaning depends on social relations, discursive uses, cultural conventions, and stylistic effects. In the case of Hindi, Sinha et al. (2005) observed that certain sociocultural meanings are grammaticalized, since honorifics can be expressed through plural pronouns and verb agreement, while echo and expressive words introduce emphatic or semantic values for which there is not always a direct correspondence in the target language. These phenomena show that pragmatic transfer does not constitute a layer independent of grammar, but can be encoded in morphological and syntactic choices.

The difficulty intensifies in literary discourse, where cultural meaning and aesthetic configuration often operate together. Tewari and Baghel (2026) found, in Hindi-English poetry, that a technically correct translation could preserve the surface meaning while simultaneously weakening the emotional impact, metaphor, or cultural value of the text. The authors found better results when the model was adapted using poetic data and stylistic cues, confirming that general translation resources alone do not adequately represent the specificities of the literary domain. This problem is also observed in creative phraseology. Comparing human translation, NMT, and LLMs on five English-Spanish literary phraseological units, Noriega-Santiáñez and Corpas Pastor (2025a) found that human translation performed best overall; NMT systems excelled particularly in morphosyntactic accuracy, while LLMs—especially ChatGPT—showed greater potential for generating creative solutions.

However, the move toward generative models introduces a tension distinct from simple cultural loss. Wang et al. (2026) found that multi-agent systems could correctly recognize certain cultural referents and yet still select or execute inappropriate strategies for translating them. In some cases, cultural information was identified during the intermediate stages of processing but did not reach the final translation correctly. This finding allows us to distinguish between the availability of cultural knowledge and the competence to convert it into a relevant translation decision.

Added to this limitation is a phenomenon that is the inverse of the traditional flattening of machine translation. Wang et al. (2026) observed that optimizing LLMs to produce stylistic effects could introduce metaphors, intensifications, or dramatic devices not present in the original. Although these interventions could improve fluency and aesthetic acceptance, they also shifted the output toward adaptation or rewriting, compromising the authorial voice and cultural specificity. Therefore, the pragmatic-cultural inadequacy of contemporary AI is not manifested solely through omission. The results allow us to identify three distinct movements: loss or flattening, when nuances of the original disappear; neutralization, when culturally or pragmatically marked forms are replaced by less specific

solutions; and generative expansion, when the system incorporates stylistic or interpretive content not supported by the source text.

This last possibility suggests using the term "cultural hallucination" with caution. In the analyzed sample, many of the problems correspond more precisely to literalization, neutralization, explication, inappropriate strategic selection, or generative addition than to hallucination in the strict sense. Even when post-editing with LLM improves expressiveness, Tewari and Baghel (2026) acknowledge the risk of unsupported reinterpretation or generation, especially in religious or culturally sensitive poetry.

For the Spanish–Hindi pair, the gap remains particularly significant. None of the 27 studies directly examined the transfer of Spanish address systems—*tú, vos, usted*—to Hindi pragmatic distinctions, nor did they experimentally evaluate how LLMs preserve register, politeness, or sociolinguistic variation in Spanish–Hindi literature. Consequently, the review identifies converging evidence regarding the vulnerability of these dimensions, but their specific behavior in this language pair remains an open empirical question.

From the rule to the instructional model: reconfiguration of mediation and post-editing

Diachronic analysis of the corpus shows that the evolution of machine translation has not only modified computational architectures but also the role of human intervention in the translation process. In the first systems included in the review, human involvement was primarily at the end of a rule-governed sequence. Sinha et al. (1995) designed ANGLABHARTI as an assisted system that generated a target pseudostructure, retained alternatives when it could not resolve certain ambiguities, and subsequently used a human post-editing package to make final corrections. This approach was maintained, albeit with greater hybridization, in AnglaHindi: Sinha and Jain (2003) combined rules, statistically derived examples from the corpus, and human post-editing, explicitly acknowledging that automatic generation alone still did not achieve a high-quality translation.

With the consolidation of statistical and neural paradigms, human intervention gradually shifted from correcting manifestly defective structures to post-editing linguistically more plausible proposals. In an experiment with six professional literary translators, Toral et al. (2018) found that post-editing with PBMT increased productivity by 18%, while post-editing with NMT increased it by 36% compared to translation from scratch. NMT also reduced keystrokes by 23%, compared to 9% with PBMT. However, the changes in cognitive effort were less linear, as post-editing produced fewer pauses, but these were longer. The same authors also pointed out that the machine's initial proposal can act as a priming element, favoring solutions close to automatic output, a particularly sensitive issue when the literary purpose involves not only preserving meaning but also reproducing a reading experience (Toral et al., 2018).

Post-editing is not necessarily the final stage of the process, as Macken et al. found in their study of professional English-Dutch literary translations. (2022) analyzed a workflow consisting of machine translation, post-editing, and revision, and found that more modifications were made during revision than during post-editing itself, and that these interventions were qualitatively different. This result demonstrates that error correction of the machine output and the final construction of a publishable literary text are partially distinct operations.

The introduction of LLMs further modifies this relationship because it allows intervention before, during, and after generation through natural language instructions. In literary translation into Spanish, Catalan, Dutch, and Chinese, Du et al. (2025) demonstrated that ChatGPT's creative performance varied according to the prompt, temperature, and text granularity. A minimal instruction explicitly aimed at creative translation, combined with a temperature of 1.0, produced the model's best results and outperformed DeepL in Spanish, Dutch, and Chinese, although it consistently fell short of human translation. The output thus ceases to be a fixed product that the translator merely accepts or corrects, as the instructions become a variable capable of deliberately modifying the translator's behavior.

This interactive logic reaches greater complexity in multi-agent systems. For example, Wang et al. (2026) compared two architectures: one that mimicked phases of professional practice and another specifically designed to meet the capabilities of LLMs, distributing strategic planning, stylistic refinement, and consistency control among specialized agents. Both workflows achieved levels of accuracy comparable to professional translations, while the system geared toward LLM capabilities obtained better stylistic and poetic language ratings, although it also introduced content foreign to the original. In poetry, Resende and Hadley (2026) also documented models capable of generating drafts, analyzing formal constraints, responding to corrective instructions, and even prompting the

user to specify whether they wish to prioritize formal fidelity, literalness, or other objectives. The authors interpret this capability as a shift from single-source generation toward an interaction closer to the negotiation of translation priorities, without eliminating the need for specialized supervision.

The studies show a trajectory that can be summarized as proofreading → post-editing → revision → instruction → selection and coordination of workflows. Therefore, the technological evolution observed in the corpus does not imply the disappearance of human mediation, but rather a reconfiguration of the translator's agency: control progressively shifts from the material production of each segment toward the formulation of objectives, the evaluation of alternatives, and the validation of final decisions. Although this pattern is highly transferable to the Spanish-Hindi problem, none of the studies analyzed empirically examines a literary flow based on an LLM for this specific linguistic direction.

The classroom as a human-AI laboratory: cognitive load and reconfiguration of translation competence

The educational and cognitive evidence from the corpus challenges the idea that the incorporation of machine translation or GenAI produces, in and of itself, a uniform reduction in translator effort. In an eye-tracking study of English-Spanish post-editing, Rojo López et al. (2024) initially compared 25 professionals and 27 students working with NMT and SMT outputs. The results showed no statistically significant differences between the two groups in the time spent or the overall duration of fixations, nor were significant time differences found between the two machine translation systems. However, inferential analysis showed that fixations were significantly longer when post-editing SMT than NMT for both professionals and students, suggesting less cognitive effort associated with neural output in this specific dimension.

This result is particularly relevant because Rojo López et al. (2024) found a virtually negligible correlation between the total time spent on the task and the duration of fixations. In other words, performing a task more quickly does not necessarily imply that it required less cognitive processing. Likewise, both professionals and students dedicated a substantial proportion of their time to external searches, demonstrating that post-editing incorporates verification and consultation activities that are not adequately reflected by purely time-based indicators. This distinction aligns with the results previously obtained by Toral et al. (2018) in professional literary translation, which showed that, although post-editing using NMT increased productivity by 36% and reduced the number of keystrokes by 23%, it produced fewer pauses, but of longer duration, than translation from scratch. Therefore, productivity, technical effort, and cognitive processing are related but not interchangeable dimensions.

The incorporation of GenAI introduces additional complexity because it shifts some of the activity toward formulating instructions, checking output, and making decisions about generated alternatives. In an eight-week training intervention with 35 university students, Mau et al. (2026) analyzed cognitive load during a GenAI-assisted legal translation task and distinguished two profiles: an Engaged but Strained group ($n = 26$), characterized by higher levels of intrinsic and extrinsic load, and a Focused and Efficient group ($n = 9$), with lower levels in both dimensions and a higher German load. Students in the second group also showed significantly greater engagement in understanding the source text, while those in the first group relied more heavily on post-editing strategies.

The results of Mau et al. (2026) also reveal that a higher intrinsic workload significantly reduced the likelihood of students actively engaging in language processing; for each unit increase in this variable, the probability of engaging in such processing decreased by approximately 82%. At the same time, the extrinsic workload associated with interacting with GenAI showed a positive (though not statistically significant) trend toward greater language engagement, presumably because verifying the automated output requires critical examination. Thus, the workload introduced by the technology cannot be conceptualized solely as interference: under certain pedagogical conditions, the effort of detecting errors, comparing solutions, and justifying corrections can become a learning opportunity.

This interpretation is reinforced by a paradox identified by Mau et al. (2026), who argue that low levels of perceived workload, when combined with excessive reliance on GenAI, can weaken metacognitive vigilance and lead students to attribute achievements derived primarily from the tool to their own learning. Therefore, the authors propose that training should not simply aim to minimize cognitive load, but rather transform the inevitable difficulties of human-AI interaction into a learning load oriented toward critical evaluation, knowledge construction, and metacognitive regulation.

Taken together, these results suggest that the instrumental competence required in AI-mediated environments transcends the operational mastery of a tool. The corpus points to a competence that integrates comprehension of

the source text, evaluation of the output, verification of information, strategic selection, justification of decisions, and metacognitive regulation. However, this conclusion must be carefully qualified, as the most direct educational evidence comes from English-Spanish post-editing and GenAI-assisted legal translation, not from Spanish-Hindi literary translation. None of the 27 studies analyzed examined a university environment specifically dedicated to AI-mediated Spanish-Hindi literary translation, so transferring these results to the literary classroom is a pedagogically sound inference, but one that is still pending empirical validation.

Voice, authorship, and agency: human control of aesthetic decision-making

The final axis of the analysis reveals that the effects of AI on literary translation cannot be evaluated solely through accuracy, fluency, or productivity, as technological mediation also affects the translator's discursive presence, the preservation of the text's voices, and the attribution of creative decisions. From an ethical perspective, Taivalkoski-Shilov (2019) argues that the quality of literary translation must be understood holistically, relating product, process, and production conditions, and identifies voice as one of the dimensions that machine translation studies have insufficiently addressed. This difficulty stems, among other factors, from the multivocality and heteroglossia characteristic of numerous literary texts, in which the stylistic configuration of narrators and characters can simultaneously fulfill aesthetic, ideological, and thematic functions.

The empirical evidence of Kenny and Winters (2020) shows that this concern is not merely theoretical. When comparing Hans-Christian Oeser's original translation with his subsequent post-editing of a fragment of the novel, the authors found that the features previously associated with his style appeared approximately one-third less frequently in the post-edited text. His textual voice remained identifiable, but it was attenuated by the initial presence of the neural input. At the same time, the 73 comments recorded during post-editing revealed a strong contextual voice, through which the translator justified lexical, stylistic, and aesthetic choices (Kenny & Winters, 2020). This result suggests that automated mediation does not necessarily eliminate translator subjectivity. However, it can shift the space in which it manifests itself, with a logic that moves from textual generation toward intervention, evaluation, and justification of decisions.

Contemporary LLMs, however, complicate this relationship because their generative capacity can expand, and not merely restrict, the creative space. Resende and Hadley (2026) show that models can be used during pre-translation to make linguistic, formal, cultural, and symbolic constraints explicit, as well as during translation to generate drafts, alternatives, or versions conditioned by specific aesthetic objectives. However, the authors maintain that the final decision rests with the translator, since it is up to them to determine whether a metrical, rhythmic, or stylistic proposal is compatible with their objectives, and the performance of models continues to vary across languages and tasks. Consequently, Resende and Hadley (2026) do not present LLMs as substitutes for creativity or human judgment, but rather as tools capable of extending them by providing access to alternatives and knowledge that the translator can incorporate, modify, or reject.

This potential expansion of creativity, however, encounters a fundamental limit in authorial fidelity. In their comparison between professional translations and multi-agent LLM systems, Wang et al. (2026) found that architectures specifically designed to leverage the model's capabilities could achieve high stylistic results and even receive higher ratings than human translation in certain aspects, especially fluency. Nevertheless, the same system occasionally introduced content not present in the original. The authors warn that this tendency toward embellishment can blur the line between translation, adaptation, and rewriting, since autonomously generated metaphors or intensifications can produce an aesthetically appealing text while simultaneously altering the author's voice, style, or intention (Wang et al., 2026).

The professional perspective gathered by Noriega-Santiáñez and Corpas Pastor (2025b) reinforces the relevance of this tension. Among literary translators surveyed in Spain, 94.1% expressed concern about the quality of work produced using GenAI, 88.2% pointed to ethical and legal problems, 80.4% expressed concern about a possible devaluation of the profession, and 84.3% about a reduction in job opportunities. Participants also warned about the potential standardization of language and the impoverishment of the literary experience. However, their positions were not homogeneous: the largest group adopted a pragmatic perspective, anticipating a selective incorporation of technology in areas where it could be useful without necessarily displacing human creativity (Noriega-Santiáñez & Corpas Pastor, 2025b).

Overall, the corpus does not support a simple opposition between human creativity and automation. The results show a tension between attenuation, amplification, and displacement of agency, largely determined by the type

of workflow and the degree of control retained by the translator. The central question thus shifts from whether AI can produce literarily convincing formulations to who defines, evaluates, and assumes responsibility for the aesthetic decisions that shape the final text. For Spanish-Hindi literary translation, this issue remains without direct empirical research within the analyzed sample, particularly regarding the simultaneous preservation of authorial voice, translator's idiolect, and intercultural polyphony.

DISCUSSION

The results of this review show that the evolution of machine translation has not eliminated the difficulties associated with linguistic, cultural, and stylistic distance, but rather has modified their location and visibility. Early systems made their limitations explicit through grammatical errors, word ordering problems, and unnatural constructions. Subsequently, NMT substantially increased fluency and reduced some of the post-editing technical effort, while LLMs incorporated capabilities for interaction, reformulation, contextualization, and creative generation. However, this evolution does not allow us to establish a linear sequence in which each paradigm definitively solves the problems of the previous one.

In literature, quality continues to depend on dimensions that go beyond propositional adequacy, such as voice, register, metaphor, cultural coherence, rhythm, and the reading experience. This interpretation aligns with Guerberof-Arenas and Toral (2022), who found greater creativity in human translations than in post-edited and machine translations, but also with Guerberof-Arenas and Toral (2024), whose reception study demonstrated that the preference for human or post-edited translations can vary between languages and reader communities. Therefore, technology-mediated literary quality should be understood as contextual and multidimensional, and not as a property reducible to fluency or similarity to a reference.

This finding is particularly relevant for Spanish–Hindi, since the review documented significant advances in multilingual translation and in the computational processing of Hindi, but also showed that specifically literary evidence for this language pair remains extremely limited. Consequently, it is not methodologically legitimate to extrapolate the positive results obtained in pairs with greater resources to conclude that structures such as the Spanish subjunctive, aspectuality, Hindi ergativity, forms of address, or literary polyphony are processed with equal efficiency.

In Hindi-English poetry, Tewari and Baghel (2026) found that domain adaptation and stylistic cues improved the translation of metaphor, tone, and emotional resonance, confirming that the overall performance of the model does not replace the need for linguistically and generically relevant resources. In this regard, recent educational literature on corpora agrees that translator training should include the ability to understand, construct, and evaluate the data that underpin the technologies used (Krüger & Hackenbuchner, 2024; Wu et al., 2025).

A first emerging conceptual finding is that the contemporary problem of literary AI can no longer be described exclusively by the traditional categories of loss, literalization, or neutralization. At the neural stage, the evidence showed a relatively clear restrictive effect: Guerberof-Arenas and Toral (2022) found that even an NMT system trained on literary information tended toward more literal and less creative solutions than those produced by human translators. Recent literature, however, introduces a different scenario. Through stylometric analysis of online Chinese literature translations, Yao et al. (2025) found that GPT-4 productions could approximate human translations in lexical, syntactic, and content features, to the point of blurring the stylistic distinction between the two modalities.

This greater aesthetic plausibility does not necessarily equate to greater literary fidelity. In the corpus of this review, Wang et al. (2026) found that multi-agent systems could achieve accuracy comparable to professionals and be preferred by evaluators in certain dimensions of fluency and style, but they could also introduce content not supported by the original. This phenomenon forces us to reconsider the very concept of error, since a system can err not only by omitting, simplifying, or neutralizing, but also by embellishing, clarifying, dramatizing, or completing beyond what is authorized by the source text. The problem then ceases to be exclusively a generation deficit and becomes also an excess of generation.

From this convergence, we propose the category of generative expansion, understood as the incorporation of semantic, cultural, or stylistic material that increases the apparent richness of the target text, but whose legitimacy cannot be sufficiently justified by the original. This category should not be automatically confused with hallucination, since some expansions may constitute explications, adaptations, or defensible creative decisions, while others

correspond to unfounded additions. The translation challenge lies precisely in determining this boundary. Therefore, the risk of contemporary literary AI can be considered bidirectional: on the one hand, loss, neutralization, and standardization; on the other, overinterpretation and generative expansion. The heterogeneity found by Guerberof-Arenas and Toral (2024) in the reception of human, machine, and post-edited translations further reinforces the idea that no technological modality can be declared universally superior, regardless of the language, the text, and the reader.

The second emerging finding stems paradoxically from the increased quality of the systems. As machine translations become grammatically fluent, contextually plausible, and stylistically convincing, the translator's difficulty shifts from producing an initial solution to determining whether a seemingly good solution is truly adequate. The contemporary problem, therefore, cannot be expressed solely by asking how much a machine can translate, but rather by asking a more demanding question focused on what skills a person needs to recognize when they should not accept what the machine proposes.

This interpretation aligns with the human-centered AI perspective developed by Jiménez-Crespo (2025), who argues that translator training should reinforce precisely those dimensions where human value depends on creativity, contextualization, narrative, common sense, judgment, and the ability to maintain control and autonomy over the system. Similarly, Knoth et al. (2024) found that greater prompt engineering competence predicts higher-quality outputs, but they linked this ability to broader AI literacy. Prompting, therefore, constitutes a new instrumental skill, but it cannot become synonymous with translation competence: generating a better response does not yet demonstrate that the user can correctly evaluate it.

This distinction becomes especially important for novice students. In their study on this topic, Zhang and Doherty (2025) found that some participants expressed confidence in their ability to identify and correct AI errors while demonstrating an insufficient understanding of its potential negative effects and ethical issues. The authors also warn that errors masked by seemingly fluent output can be difficult to recognize and advocate for integrating AI literacy from the outset of translator training. Fluency, in this sense, can become an epistemological challenge, since the less a translation appears to need correction, the greater the need for a competent reader capable of questioning it.

The results regarding cognitive load point in the same direction. In the corpus analyzed, Mau et al. (2026) showed that interaction with GenAI generates differentiated cognitive profiles and that high confidence combined with a low perceived load can weaken metacognitive vigilance. More recent literature also does not allow us to assume that adding AI uniformly reduces effort, as pedagogical strategies do matter. Tian et al. (2025) found that requiring students to formulate their own interpretation before consulting GenAI reduced overall effort and, simultaneously, increased higher-order thinking compared to a more passive use of the tool. The conceptual consequence is relevant, as it means that a lower workload does not necessarily equate to better learning, and greater difficulty is not necessarily pedagogically detrimental; what is decisive is which cognitive processes are being activated.

These transformations have a direct implication for higher education. If producing a linguistically acceptable version ceases to be the defining characteristic of translation work, university programs must avoid limiting technological instruction to the operational handling of platforms. Ma et al. (2024) conceptualized ChatGPT literacy through dimensions that include knowledge of possibilities and limitations, formulation of instructions, evaluation of responses, educational assessment, and ethics. In a more specifically translation-related field, Krüger and Hackenbuchner (2024) propose that professional machine translation literacy is also associated with data comprehension and management. These approaches suggest that contemporary technological competence must include knowing what a system produces, what limitations it has, what evidence can be relied upon, and under what conditions its output should be rejected.

From this perspective, the results allow us to propose a human-AI translation competence structured around five interdependent operations: deep reading of the source text, critical literacy regarding technologies and their data, strategic design of the interaction, evaluation and verification of outputs, and metacognitive and ethical regulation of the final decision. This formulation aligns with Cook et al. (2025), who argue that literary translation education should focus on the deep construction of meaning and human added value, and with Jiménez-Crespo (2025), for whom curriculum redesign must preserve the centrality of human control. Therefore, translator training should not compete with AI in terms of speed of generation, but rather develop the capacity to produce decisions that can be argued linguistically, culturally, aesthetically, and ethically.

This conception also modifies the design of classroom activities, as demonstrated by the evidence from Łoboda and Mastela (2023). These authors indicate that post-editing culturally marked texts can function simultaneously as translator training and as a critical evaluation of the technology. Wu et al. (2025) extend this logic by having students construct corpora themselves and using text mining to formulate retranslation plans, while Tian et al. (2025) demonstrate that structuring the interaction so that students first develop their own hypotheses reduces passive dependence on GenAI. For Spanish-Hindi, this evidence suggests a classroom organized as a comparative laboratory, where problematic phenomena (subjunctive mood, aspect, forms of address, metaphors, discourse markers, or cultural references) can be confronted using human versions and multiple AI outputs, without presupposing that any of them automatically constitutes the correct answer.

Instead of evaluating only the finished product, it would be necessary to observe what errors the student detects, which ones they omit, what modifications they introduce, and how they justify their choices. Bodart et al. (2024) developed precisely a pedagogical taxonomy that distinguishes successful, unnecessary, incomplete, failed, or absent interventions in post-editing, while Diels et al. (2025) found that a corpus-based intervention could modify revision processes even when its effects were not statistically significant on the final product. Similarly, Guerberof-Arenas et al. (2024) did not find sufficient quantitative evidence to conclude that post-editing training directly increased creativity, although students subsequently showed a greater willingness to make creative changes and greater confidence when faced with working memory problems. These results are especially relevant to educational contexts: learning is not always immediately manifested in a superior final translation, but also in the transformation of the process by which the student makes decisions.

AI can also expand feedback possibilities, but evidence suggests maintaining a complementary model. Su et al. (2025) found that ChatGPT tended to offer more direct solutions and general observations, while teacher feedback was more elaborate and received greater student acceptance. In contrast, using a pipeline specifically designed for translator assessment, Jiao et al. (2025) obtained LLM feedback with high agreement with the expert and a positive perception among students. This apparent contradiction reaffirms a recurring thesis of this review: performance does not simply depend on whether or not AI is used, but on how the human-machine flow is configured and what pedagogical function is assigned to it.

Consequently, the educational incorporation of AI should avoid both its uncritical prohibition and its instrumentalist adoption. Moorhouse et al. (2024) demonstrated that professional competence in integrating GenAI can be developed through explicit training, while Freeth and Toto (2026) found that an AI-in-the-loop pedagogy can lead to a more critical and selective use of the technology. From this perspective, the main educational objective would not be to get students to use AI more, but rather to ensure they know when to use it, what to use it for, how to question it, how to verify it, and when to move away from it.

Based on these findings, AI can be conceptualized in the teaching of literary translation as a heuristic mediator under human governance. Its pedagogical value lies less in providing the translation itself than in expanding the space of possibilities that students must consider: generating alternatives, revealing contrasts, making decisions visible, provoking disagreements, and requiring them to justify their preferences. In the Spanish-Hindi case, where this review identifies a clear lack of direct literary evidence, this function is especially relevant: the absence of established solutions can become a formative scenario for developing contrastive competence, critical evaluation, and intercultural awareness. Thus, the shift from flattening to generative expansion and from generation to judgment converges on the same educational consequence: the more capable AI is of producing plausible texts, the more necessary it will be to train translators capable of exercising a critical, informed, and justifiable agency over them.

CONCLUSIONS

The review concludes that the study of AI-mediated literary translation requires moving beyond evaluations focused solely on the technical performance of the systems. In literary texts, quality must be analyzed from a multidimensional perspective that integrates linguistic appropriateness, pragmatic-cultural coherence, stylistic preservation, human intervention, and conditions of use. This approach is especially necessary when working with typologically distant languages and unequal resource availability, as the general progress of the models does not provide sufficient evidence of their performance in specific language combinations.

For Spanish-Hindi, the main conclusion is epistemological. There is a sufficient basis to justify the relevance of the problem, but not yet enough empirical evidence to formulate generalizations about the performance of contemporary systems in literary translation between the two languages. The field therefore demands research

specifically designed for this pair, capable of isolating morphosyntactic, pragmatic, and stylistic phenomena and examining them in real literary contexts. The construction of specialized parallel corpora, annotated not only linguistically but also in pragmatic and stylistic dimensions, is a priority for moving toward more consistent evaluations.

In the educational field, this review leads to a rethinking of the purpose of incorporating AI into university translator training. Its integration is academically relevant when it allows for the observable decision-making process, the comparison of alternatives, and the subjection of technological solutions to linguistic, cultural, and aesthetic argumentation. Consequently, curricular and evaluative designs should place greater emphasis on the traceability of decisions, the ability to justify modifications, and reflection on the translation process, preventing the quality of learning from being inferred solely from the final product.

Future research should advance in four complementary directions: experimental evaluation of Spanish–Hindi literary translation using NMT and LLM models; development and validation of specialized literary corpora and benchmarks; longitudinal studies on university training and the evolution of translation competence in human–AI environments; and comparative analyses of different interaction flows, including multi-agent systems, post-editing, guided generation, and human-assisted translation. It will also be necessary to more systematically incorporate reader reception, the variation between literary genres, and the authorial, professional, and educational consequences of these new processes. This will allow research to shift from describing technological capabilities to constructing responsible and pedagogically sound models of AI-assisted literary translation.

REFERENCES

Al-Batineh, M., & Al Tenaijy, M. (2024). Adapting to technological change: An investigation of translator training and the translation market in the Arab world. *Heliyon*, *10*(7), e28535. https://doi.org/10.1016/j.heliyon.2024.e28535

Bhattacharjee, S., Gain, B., & Ekbal, A. (2024). Domain Dynamics: Evaluating Large Language Models in English-Hindi Translation. En B. Haddow, T. Kocmi, P. Koehn, & C. Monz (Eds.), *Proceedings of the Ninth Conference on Machine Translation* (pp. 341–354). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.wmt-1.27

Bodart, R., Piette, J., & Lefer, M.-A. (2024). The *Machine Translation Post-Editing Annotation System* (MTPEAS): A standardized and user-friendly taxonomy for student post-editing quality assessment. *Translation Spaces*, *13*(2), 265–292. https://doi.org/10.1075/ts.24002.bod

Bojar, O., Straňák, P., & Zeman, D. (2010). Data Issues in English-to-Hindi Machine Translation. En N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, & D. Tapias (Eds.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. European Language Resources Association (ELRA). https://aclanthology.org/L10-1524/

Cook, A., Knaggs, A., & Sun, J. (2025). Exploring the goals of translation education: Adding human value to literary translation through deep meaning making and artistry. *The Interpreter and Translator Trainer*, *19*(3–4), 254–276. https://doi.org/10.1080/1750399X.2025.2543208

Dave, S., Parikh, J., & Bhattacharyya, P. (2001). Interlingua-based English–Hindi Machine Translation and Language Divergence. *Machine Translation*, *16*(4), 251–304. https://doi.org/10.1023/A:1021902704523

Diels, E., Ureel, J. J. J., Robert, I. S., & Strobl, C. (2025). The effects of corpus-focused instruction on the development of stylistic translation revision competence in future translators. *Translation Spaces*, *14*(2), 221–252. https://doi.org/10.1075/ts.24038.die

Du, S., Arenas, A. G., Toral, A., Gerrits, K., & Borillo, J. M. (2025). Optimising ChatGPT for creativity in literary translation: A case study from English into Dutch, Chinese, Catalan and Spanish. En P. Bouillon, J. Gerlach, S. Girletti, L. Volkart, R. Rubino, R. Sennrich, A. C. Farinha, M. Gaido, J. Daems, D. Kenny, H. Moniz, & S. Szoc (Eds.), *Proceedings of Machine Translation Summit XX: Volume 1* (pp. 578–591). European Association for Machine Translation. https://aclanthology.org/2025.mtsummit-1.44/

Ehrensberger-Dow, M., Delorme Benites, A., & Lehr, C. (2023). A new role for translators and trainers: MT literacy consultants. *The Interpreter and Translator Trainer*, *17*(3), 393–411. https://doi.org/10.1080/1750399X.2023.2237328

Fan, A., Bhosale, S., Schwenk, H., Ma, Z., El-Kishky, A., Goyal, S., Baines, M., Celebi, O., Wenzek, G., Chaudhary, V., Goyal, N., Birch, T., Liptchinsky, V., Edunov, S., Auli, M., & Joulin, A. (2021). Beyond English-Centric Multilingual Machine Translation. *Journal of Machine Learning Research*, *22*(107), 1–48. http://jmlr.org/papers/v22/20-1307.html

Freeth, P. J., & Toto, P. (2026). The impact of AI-in-the-loop pedagogy: An action research approach to integrating AI technologies in translator training. *The Interpreter and Translator Trainer*, 1–18. https://doi.org/10.1080/1750399X.2026.2646863

González-Pastor, D. (2024). La traducción automática y la formación de traductores en España: Perspectivas desde la industria y el ámbito académico. *Mutatis Mutandis. Revista Latinoamericana de Traducción*, *17*(1). https://doi.org/10.17533/udea.mut.v17n1a06

Goyal, N., Gao, C., Chaudhary, V., Chen, P.-J., Wenzek, G., Ju, D., Krishnan, S., Ranzato, M., Guzmán, F., & Fan, A. (2022). The FLORES-101 Evaluation Benchmark for Low-Resource and Multilingual Machine Translation. *Transactions of the Association for Computational Linguistics*, *10*, 522–538. https://doi.org/10.1162/tacl_a_00474

Goyal, V., Mishra, P., & Sharma, D. M. (2020). Linguistically Informed Hindi-English Neural Machine Translation. En N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 3698–3703). European Language Resources Association. https://aclanthology.org/2020.lrec-1.456/

Guerberof-Arenas, A., & Toral, A. (2022). Creativity in translation: Machine translation as a constraint for literary texts. *Translation Spaces*, *11*(2), 184–212. https://doi.org/10.1075/ts.21025.gue

Guerberof-Arenas, A., & Toral, A. (2024). To be or not to be: A translation reception study of a literary text translated into Dutch and Catalan using machine translation. *Target. International Journal of Translation Studies*, *36*(2), 215–244. https://doi.org/10.1075/target.22134.gue

Guerberof-Arenas, A., Valdez, S., & Dorst, A. G. (2024). Does training in post-editing affect creativity? *The Journal of Specialised Translation*, (41), 74–97. https://doi.org/10.26034/cm.jostrans.2024.4712

He, S., Hao, Y., Liu, S., Liu, H., & Li, H. (2022). Research on translation technology teaching in Chinese publications and in international English-language publications (1999-2020): A bibliometric analysis. *The Interpreter and Translator Trainer*, *16*(3), 275–293. https://doi.org/10.1080/1750399X.2022.2101848

He, Y., & Tao, Y. (2022). Unity of knowing and acting: An empirical study on a curriculum approach to developing students' translation technological thinking competence. *The Interpreter and Translator Trainer*, *16*(3), 348–366. https://doi.org/10.1080/1750399X.2022.2101849

Huang, Y., & Cheung, A. K. F. (2026). Exploring AI's performance in literary autobiography translation: How closely do AI models match human translation. *Humanities and Social Sciences Communications*, *13*(1), 518. https://doi.org/10.1057/s41599-026-06630-4

Ibáñez Moreno, A., & Domínguez Mora, M. E. (2025). Google Translate versus DeepL in Spanish to English translation of *Don* Quixote. *Translation and Translanguaging in Multilingual Contexts*, *11*(1), 65–87. https://doi.org/10.1075/ttmc.00154.iba

Jiao, H., Hu, W., & Zhang, X. (2025). *To eat or to feed*: Can large language models provide useful feedback in translation education? *The Interpreter and Translator Trainer*, *19*(3–4), 317–337. https://doi.org/10.1080/1750399X.2025.2533074

Jiménez-Crespo, M. A. (2025). "If students translate like a robot ... " or how research on human-centered AI and intelligence augmentation can help realign translation education. *The Interpreter and Translator Trainer*, *19*(3–4), 277–295. https://doi.org/10.1080/1750399X.2025.2542022

Karpinska, M., & Iyyer, M. (2023). Large Language Models Effectively Leverage Document-level Context for Literary Translation, but Critical Errors Persist. *Proceedings of the Eighth Conference on Machine Translation*, 419–451. https://doi.org/10.18653/v1/2023.wmt-1.41

Kenny, D., & Winters, M. (2020). Machine translation, ethics and the literary translator's voice. *Translation Spaces*, *9*(1), 123–149. https://doi.org/10.1075/ts.00024.ken

Knoth, N., Tolzin, A., Janson, A., & Leimeister, J. M. (2024). AI literacy and its implications for prompt engineering strategies. *Computers and Education: Artificial Intelligence*, *6*, 100225. https://doi.org/10.1016/j.caeai.2024.100225

Krüger, R., & Hackenbuchner, J. (2024). A competence matrix for machine translation-oriented data literacy teaching. *Target. International Journal of Translation Studies*, *36*(2), 245–275. https://doi.org/10.1075/target.22127.kru

Kwok, H. L., Shi, Y., Xu, H., Li, D., & Liu, K. (2025). GenAI as a translation assistant? A corpus-based study on lexical and syntactic complexity of GPT-post-edited learner translation. *System*, *130*, 103618. https://doi.org/10.1016/j.system.2025.103618

Łoboda, K., & Mastela, O. (2023). Machine translation and culture-bound texts in translator education: A pilot study. *The Interpreter and Translator Trainer*, *17*(3), 503–525. https://doi.org/10.1080/1750399X.2023.2238328

Ma, Q., Crosthwaite, P., Sun, D., & Zou, D. (2024). Exploring ChatGPT literacy in language education: A global perspective and comprehensive approach. *Computers and Education: Artificial Intelligence*, *7*, 100278. https://doi.org/10.1016/j.caeai.2024.100278

Macken, L., Vanroy, B., Desmet, L., & Tezcan, A. (2022). Literary translation as a three-stage process: Machine translation, post-editing and revision. En H. Moniz, L. Macken, A. Rufener, L. Barrault, M. R. Costa-jussà, C. Declercq, M. Koponen, E. Kemp, S. Pilos, M. L. Forcada, C. Scarton, J. Van den Bogaert, J. Daems, A. Tezcan, B. Vanroy, & M. Fonteyne (Eds.), *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation* (pp. 101–110). European Association for Machine Translation. https://aclanthology.org/2022.eamt-1.13/

Manakhimova, S., Avramidis, E., Macketanz, V., Lapshinova-Koltunski, E., Bagdasarov, S., & Möller, S. (2023). Linguistically Motivated Evaluation of the 2023 State-of-the-art Machine Translation: Can ChatGPT Outperform NMT? *Proceedings of the Eighth Conference on Machine Translation*, 224–245. https://doi.org/10.18653/v1/2023.wmt-1.23

Mau, B.-R., Wu, Y.-P., & Feng, H.-H. (2026). Mind the cognitive load gap: Student translators' cognitive demands and translation behaviors in GenAI-assisted legal translation. *System*, *139*, 104033. https://doi.org/10.1016/j.system.2026.104033

Moorhouse, B. L., Wan, Y., Wu, C., Kohnke, L., Ho, T. Y., & Kwong, T. (2024). Developing language teachers' professional generative AI competence: An intervention study in an initial language teacher education course. *System*, *125*, 103399. https://doi.org/10.1016/j.system.2024.103399

NLLB Team. (2024). Scaling neural machine translation to 200 languages. *Nature*, *630*(8018), 841–846. https://doi.org/10.1038/s41586-024-07335-x

Noriega-Santiáñez, L., & Corpas Pastor, G. (2025a). Measuring Creative Phraseology in Literature: Machine Translation Systems Versus Large Language Models. *Yearbook of Phraseology*, *16*(1), 125–152. https://doi.org/10.1515/phras-2025-0006

Noriega-Santiáñez, L., & Corpas Pastor, G. (2025b). Technology and GenAI adoption among literary translators in Spain: A survey study on uses, perceptions and attitudes. *Tradumàtica tecnologies de la traducció*, (23), 137–164. https://doi.org/10.5565/rev/tradumatica.505

Peng, K., Ding, L., Zhong, Q., Shen, L., Liu, X., Zhang, M., Ouyang, Y., & Tao, D. (2023). Towards Making the Most of ChatGPT for Machine Translation. *Findings of the Association for Computational Linguistics: EMNLP 2023*, 5622–5633. https://doi.org/10.18653/v1/2023.findings-emnlp.373

Prieto Ramos, F. (2024). Revisiting translator competence in the age of artificial intelligence: The case of legal and institutional translation. *The Interpreter and Translator Trainer*, *18*(2), 148–173. https://doi.org/10.1080/1750399X.2024.2344942

Rajpoot, P., Bhat, N., & Shrivastava, A. (2024). Multimodal Machine Translation for Low-Resource Indic Languages: A Chain-of-Thought Approach Using Large Language Models. *Proceedings of the Ninth Conference on Machine Translation*, 833–838. https://doi.org/10.18653/v1/2024.wmt-1.79

Resende, N., & Hadley, J. (2026). Extending Creativity: Large Language Models and the Practice of Poetry Translation. En D. Shterionov, E. Vanmassenhove, M. De Sisto, F. Blain, J. Pourmostafa Roshan Sharami, L. Lepp, C. Manna, A. A. Rescigno, A. Karakanta, A. Rigouts Terryn, M. Lardelli, N. Resende, E. Murgolo, J. Hackenbuchner, A. Zaretskaya, M. Esplà-Gomis, T. Etchegoyhen, D. Gromann, R. Bawden, ... H. Moniz (Eds.), *Proceedings of the 26th Annual Conference of the European Association for Machine Translation (Volume 1)* (pp. 787–799). European Association for Machine Translation. https://aclanthology.org/2026.eamt-1.50/

Robinson, N., Ogayo, P., Mortensen, D. R., & Neubig, G. (2023). ChatGPT MT: Competitive for High- (but Not Low-) Resource Languages. *Proceedings of the Eighth Conference on Machine Translation*, 392–418. https://doi.org/10.18653/v1/2023.wmt-1.40

Rojo López, A. M., Vicente López, M. I., & Hvelplund, K. T. (2024). Measuring cognitive effort in post-editing: An eye-tracking study comparing professional and student translators. *Tradumàtica tecnologies de la traducció*, (22), 112–135. https://doi.org/10.5565/rev/tradumatica.418

Sinha, K., Mahesh, R., & Thakur, A. (2005). Translation divergence in English-Hindi MT. *Proceedings of the 10th EAMT Conference: Practical applications of machine translation*. https://aclanthology.org/2005.eamt-1.33/

Sinha, R. M. K., & Jain, A. (2003). AnglaHindi: An English to Hindi machine-aided translation system. *Proceedings of Machine Translation Summit IX: System Presentations*. https://aclanthology.org/2003.mtsummit-systems.15/

Sinha, R. M. K., Sivaraman, K., Agrawal, A., Jain, R., Srivastava, R., & Jain, A. (1995). ANGLABHARTI: A multilingual machine aided translation project on translation from English to Indian languages. *1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century*, *2*, 1609–1614. https://doi.org/10.1109/ICSMC.1995.538002

Sinha, R. M. K., & Thakur, A. (2005). Handling ki in Hindi for Hindi-English MT. *Proceedings of Machine Translation Summit X: Posters*, 356–353. https://aclanthology.org/2005.mtsummit-posters.6/

Su, Y., Xu, S., & Liu, K. (2025). Adapt or adopt? Examining the efficacy of ChatGPT in providing translation feedback. *The Interpreter and Translator Trainer*, *19*(3–4), 296–316. https://doi.org/10.1080/1750399X.2025.2541486

Taivalkoski-Shilov, K. (2019). Ethical issues regarding machine(-assisted) translation of literary texts. *Perspectives*, *27*(5), 689–703. https://doi.org/10.1080/0907676X.2018.1520907

Tewari, P., & Baghel, A. S. (2026). STYLISTICALLY-AWARE HINDI-ENGLISH POETIC TRANSLATION WITH MBART AND LLM-BASED POST- EDITING. *Journal of Theoretical and Applied Information Technology*, *104*(3). https://www.jatit.org/volumes/Vol104No3/36Vol104No3.pdf

Tian, S., Wang, D., Wang, J., & Zhong, W. (2025). Empowering GenAI with a guidance-based approach in MTPE learning: Effect on student translators' cognitive process, final translation quality and learning motivation. *The Interpreter and Translator Trainer*, *19*(3–4), 379–404. https://doi.org/10.1080/1750399X.2025.2534269

Tian, X. (2024). Personalized translator training in the era of digital intelligence: Opportunities, challenges, and prospects. *Heliyon*, *10*(20), e39354. https://doi.org/10.1016/j.heliyon.2024.e39354

Toral, A., & Way, A. (2015). Machine-assisted translation of literary text: A case study. *Translation Spaces*, *4*(2), 240–267. https://doi.org/10.1075/ts.4.2.04tor

Toral, A., & Way, A. (2018). What Level of Quality Can Neural Machine Translation Attain on Literary Text? En J. Moorkens, S. Castilho, F. Gaspari, & S. Doherty (Eds.), *Translation Quality Assessment* (Vol. 1, pp. 263–287). Springer International Publishing. https://doi.org/10.1007/978-3-319-91241-7_12

Toral, A., Wieling, M., & Way, A. (2018). Post-editing Effort of a Novel With Statistical and Neural Machine Translation. *Frontiers in Digital Humanities*, *5*, 9. https://doi.org/10.3389/fdigh.2018.00009

Wang, L., Lyu, C., Ji, T., Zhang, Z., Yu, D., Shi, S., & Tu, Z. (2023). Document-Level Machine Translation with Large Language Models. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 16646–16661. https://doi.org/10.18653/v1/2023.emnlp-main.1036

Wang, L., Sun, S., Wang, X., Gu, J., & Liu, K. (2026). Workflow matters: Comparing human translators and multi-agent LLMs in literary translation. *Target. International Journal of Translation Studies*. https://doi.org/10.1075/target.25081.wan

Wu, Y.-P., Feng, H.-H., & Mau, B.-R. (2025). Didactic potential of working with DIY corpora and text mining approaches in literary translation training. *The Interpreter and Translator Trainer*, *19*(2), 170–196. https://doi.org/10.1080/1750399X.2025.2488714

Yao, X., Kang, Y.-B., & McCosker, A. (2025). Missing the human touch?: A computational stylometric analysis of GPT-4 translations of online Chinese literature. *Translation Spaces*, *14*(2), 303–330. https://doi.org/10.1075/ts.24043.yao

Zerva, C., Blain, F., C. De Souza, J. G., Kanojia, D., Deoghare, S., Guerreiro, N. M., Attanasio, G., Rei, R., Orasan, C., Negri, M., Turchi, M., Chatterjee, R., Bhattacharyya, P., Freitag, M., & Martins, A. (2024). Findings of the Quality Estimation Shared Task at WMT 2024: Are LLMs Closing the Gap in QE? *Proceedings of the Ninth Conference on Machine Translation*, 82–109. https://doi.org/10.18653/v1/2024.wmt-1.3

Zhang, J., & Doherty, S. (2025). Investigating novice translation students' AI literacy in translation education. *The Interpreter and Translator Trainer*, *19*(3–4), 234–253. https://doi.org/10.1080/1750399X.2025.2541478

FINANCING
None.

CONFLICT OF INTEREST STATEMENT
None.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE
No artificial intelligence was used in the development of the article.

AUTHORSHIP CONTRIBUTION
Conceptualization: Sabyasachi Mishra.
Data curation: Sabyasachi Mishra.
Formal analysis: Sabyasachi Mishra.
Research: Sabyasachi Mishra.
Methodology: Sabyasachi Mishra.
Software: Sabyasachi Mishra.
Supervision: Sabyasachi Mishra.

Validation: Sabyasachi Mishra.
Visualization: Sabyasachi Mishra.
Writing – original draft: Sabyasachi Mishra.
Writing – proofreading and editing: Sabyasachi Mishra.