ELTR Journal, e-ISSN 2579-8235, Vol. 10, No. 1, January 2026, pp. 1-17

English Language Teaching and Research Journal
http://apspbi.or.id/eltr
English Language Education Study Program Association, Indonesia

TRANSLATION STRATEGIES IN EFL/ESL AND MTGEN/AI POST-EDITING

Lusinda Juliani[1]*, Nailah Zubaidah[2], Cyntya Isabel[3] and Refika Andriani[4]
[1,2,3,4] Lancang Kuning University,Indonesia
lusinainggolan932@gmail.com[1], refika@unilak.ac.id[4]
*correspondence: lusinainggolan932@gmail.com
https://doi.org/10.37147/eltr.v10i1.294
received 31 October 2025; accepted 27 January 2026

Abstract
This study investigates the interplay of translation techniques, machine translation (MT), and generative artificial intelligence (GenAI) within English as a Foreign Language (EFL) and English as a Second Language (ESL) settings. Utilizing a Systematic Literature Review (SLR) of 23 peer-reviewed studies from 2021–2025, it delineates prevailing methodologies in literary, business, and cultural texts, investigates determinants affecting strategy selection, and contrasts human translation with MT/GenAI post editing regarding accuracy, fluency, and cultural subtleties. Following PRISMA 2020 criteria, the analysis used descriptive statistics and thematic categorization based on both traditional and postcolonial frameworks. The results show that more and more people want a hybrid approach to translating education and practice, where MT/GenAI makes drafts and human post editing improves the quality of language and culture. The "draft by machine, craft by human" paradigm improves translation skills by connecting tactics like explicitation, compensation, and idiomatic adaptation to better readability and coherence. The study suggests combining MT/GenAI with rubric based post editing to improve translators' tech and strategic skills in modern translation teaching.

Keywords: machine translation (MT), post-editing, translation strategies

Introduction
With the increasing integration of machine translation and generative models (MT/GenAI) into professional and academic practice across EFL/ESL contexts, translation studies are entering a new phase. A total of 23 documents met our inclusion criteria (peer-reviewed/proceedings, in English/Indonesian, focused on translation strategies and/or MT/GenAI) based on an Excel corpus systematically compiled using an SLR protocol (2021–2025) from indexed databases (e.g., Scopus, Web of Science, ERIC) and complementary searches (Google Scholar). Geographically, the subjects span Europe (the Netherlands), Saudi Arabia, China, Indonesia, and Morocco, enabling cross system and cross-cultural comparisons. Through careful open synthesis, this research is crucial for updating knowledge about how traditional translation techniques interact/compete with MT/GenAI-

based post editing workflows in classrooms and workplaces (Krüger, 2021; Maid & Azmi, 2025; Yang, Aniceto, Allen, Bulusu, & Tsang, 2022).

The findings show that technology assisted translation (CAT tools, MT, LLMs) stands out (11/23), while studies on translation classes (22/23) and strong connections with cultural terms (10/23) dominate the corpus at the level of practice. These figures derive from a multistage screening process (identification, screening, eligibility, inclusion) following PRISMA 2020 with two independent raters and consensus. From post editing workflows on nonliterary (business) texts to classroom aids and comparisons of ChatGPT/Google Translate for literary works, the pattern confirms that teaching and learning are now closely tied to the digital ecosystem (Abdelhalim, Alsahil, & Alsuhaibani, 2025; Junaeny & Nirdayanti, 2023; Wang, 2021).

Preliminary results underscore the urgency of mapping the techniques actually used by students, novice translators, and experienced translators, and of examining how text types and cultural contexts influence the choice of approach within an SLR evidence framework. In classroom and business studies, for example, improvements in translation performance are linked to task design, feedback, and the introduction of MT+post editing workflows (Guo, 2023); Kilaton, 2024; Maid & Azmi, 2025). For Chinese Indonesian cultural terms, strategy preferences likewise follow local cultural sensitivities and communicative purposes (Junaeny & Nirdayanti, 2023).

This study has four objectives: (1) to identify the most common translation strategies for literary texts, business texts, and cultural terms; (2) to examine factors influencing strategy choice competence, text type, cultural context, and the availability of MT/GenAI; (3) to compare the quality of human strategies vs MT/GenAI + post-editing (accuracy, fluency, and preservation of cultural nuance); and (4) to assess how well the practice and teaching of strategies improve students' translation performance. These objectives are pursued through a synthesis plan (descriptive frequency & thematic coding), a priori inclusion–exclusion criteria, and a documented protocol (structured research questions, keyword strings, year range, country domains) to produce evidence-based pedagogical recommendations (Kilaton, 2024; Krüger, 2021; Yang, Aniceto, Allen, Bulusu, & Tsang, 2022)

The theoretical framework builds on the foundations of strategy and quality assessment that have been used and updated in the 2021–2025 corpus, including a taxonomy of procedures/strategies applied to classroom, business, and cultural-term data, as well as a quality assessment matrix weighing pragmatic/textual dimensions. In our corpus, this is operationalized through indicators of strategy (literal translation, adaptation, compensation, explicitation/implicitation) and quality benchmarks (accuracy, fluency, cultural equivalence) in recent studies on CAT/MT/LLMs and translation pedagogy (Junaeny & Nirdayanti, 2023; Maid & Azmi, 2025; Wang, 2021; Yang, 2022).

On the technology side, possibilities and limitations are drawn from studies on MT, corpora, and state of the art statistical/neural frameworks in EFL/ESL and applied education including implementations of CAT/MT in the classroom, Python-based MT teaching repositories, and computationally assisted evaluation and post editing (Akhan & Uzun, 2025; Krüger, 2021; Liu, Kwok, Liu, & Cheung, 2022; Yang, 2022). Machine learning based evaluation studies for learner translation quality also emerge in the corpus (Alkhbeer, 2025).

The socio historical context is equally important: recent literary and cultural scholarship in the corpus shows how representations of identity/minorities and power relations influence strategy choices (domestication/foreignization, adaptation, compensation) and decisions to retain/tame cultural elements for target readers. This perspective is operationalized at the data extraction stage as contextual variables (period, power relations, cultural orientation) and at the appraisal stage as potential sources of interpretive bias (A.R., 2020; Junaeny & Nirdayanti, 2023).

Sources of analysis include journal articles, conference proceedings, and reports/theses recorded in the Excel corpus. From each study we extracted data on aims, design/methods, participants/corpora, tools/technology, translation techniques, quality metrics, key findings, and limitations via a standardized extraction sheet. The diversity of regions (Saudi Arabia, China, Indonesia, Morocco, the Netherlands) and text types (literature, business, cultural terms) enables triangulation and systematic mapping of the impact of MT/GenAI and post-editing in both education and professional practice (Junaeny & Nirdayanti, 2023; Maid & Azmi, 2025; Wang, 2021).

Our study's specific foci include Chinese Indonesian cultural terminology, business texts (Business English), literary texts (including a comparison of ChatGPT vs Google Translate), translation classes (EFL/ESL students), and technology-assisted learning (CAT tools, MT, LLMs) with post editing. With a 2021–2025 time window, we frame the scope as a modified Population-Intervention Outcome Context (PICO/POCOK) domain for applied humanities SLRs where, although literary/business studies are fewer than classroom/cultural-term studies, they provide crucial units of analysis for accuracy, fluency, and nuance, especially when MT/GenAI outputs are tested through post-editing grounded in traditional techniques (A.R., 2020; Maid & Azmi, 2025; Wang, 2021).

This research contributes in two ways. Through transparent evidence mapping and thematic synthesis, we connect traditional approaches with modern practices (MT/GenAI + post-editing) in EFL/ESL contexts. Competency targets, task design, rubrics that account for cultural nuance, and practical workflows for tech-native classrooms can be integrated into curricula based on pedagogical comparisons of human vs machine + post-editing, evaluated using study quality appraisal and sensitivity to risk of bias as practiced in the studies in our corpus (Guo, 2023; Kilaton, 2024; Maid & Azmi, 2025).

Human interaction & patterns of intercivilizational communication

From manuscripts to digital platforms, linguistic hegemony shapes "standards" that in turn guide strategic choices when to be literal or to adapt, how to apply compensation and explicitation/implicitation for cultural realia. In the 2021–2025 corpus, these issues surface when MT/GenAI outputs are compared with human translations via post editing to preserve accuracy, fluency, and cultural nuance, particularly in classrooms and business texts (Maid & Azmi, 2025; Wang, 2021; Yang, 2022).

Postcolonial perspectives & their application to literary analysis

In literary translation, the ethics of domestication vs foreignization together with considerations of equivalence and systemic norms guides whether cultural elements are retained or "tamed" for target readers. In our corpus, this appears in

comparisons of human vs MT/GenAI + post editing regarding accuracy, fluency, and preservation of nuance, as well as in pedagogical implications for combining traditional techniques with technology-based workflows (Akhan & Uzun, 2025; A.R., 2020; Junaeny & Nirdayanti, 2023; Liu, 2022).

Method

Data sources & context

This study synthesizes literature on translation techniques and/or MT/GenAI in EFL/ESL contexts through a Systematic Literature Review (SLR). The main source is an internal Excel corpus that includes 23 publication records from 2021–2025 that were methodically gathered from indexed databases (Scopus, Web of Science, and ERIC) and supplementary searches (Google Scholar). Cross-system and cross-cultural comparison is made possible by the locations, which include Europe (the Netherlands), Saudi Arabia, China, Indonesia, and Morocco. Only English and Indonesian are used in publications, and the emphasis is on post-editing, CAT/MT/LLMs, and translation techniques in professional and classroom settings that fit the scope stated in the introduction. In today's translation learning and work, this setting was selected to trace the interactions between conventional methodologies and modern practice (MT/GenAI+post-editing).

Procedures & acquisition

Identification, screening, eligibility, and inclusion are the four phases of the SLR's PRISMA 2020 methodology. The search phrases (e.g., translation strategy/procedure, machine translation, post editing, LLM/GenAI, EFL/ESL, and text domains including literary, business, and cultural terms) were taken from the research topics and theoretical framework. Peer-reviewed journal articles, proceedings, or applied research in English or Indonesian, with an emphasis on translation techniques and/or MT/GenAI, years 2021–2025, EFL/ESL contexts, or pertinent professional practice were the requirements for inclusion. Opinion pieces and non-systematic reviews, publications outside of translation research, studies conducted prior to 2021, and manuscripts with unclear data or methodology were also excluded. After eliminating duplicates, two independent reviewers used a consensus approach to screen titles, abstracts, and full texts; differences were settled by discussion and third-party adjudication. The title, objectives, design, participants/corpus, tools/technology, translation methods, quality metrics, conclusions, and limitations of every included study were taken out and put onto a standardized sheet that matched the Excel column format.

Preprocessing & analysis

The data were cleaned by harmonizing technology labels (CAT/MT/LLM), standardizing terminology (strategy taxonomy: Nida, Newmark, Toury, and House as a priori categories), and standardizing entries (author/year, venue, nation, and text domain). Documents that were unclear, irrelevant, or multidomain were marked for further screening. The frequencies of strategies, text domains, and technology adoption were calculated through quantitative-descriptive analyses; the determinants of strategy choice (competence, text type, cultural context, and availability of MT/GenAI) were mapped using deductive inductive thematic coding in qualitative analyses, which included a postcolonial lens for cultural realities. For

aggregate and summary statistics, Python 3.11 (pandas 2.x, numpy 1.26.x) was utilized in the computation; Cohen's κ was used to evaluate coding reliability, with a target of ≥0.75. Limited logistic regression examined relationships (predictors: text domain/technology; outcome: acceptability/accuracy) when data allowed, presenting odds ratios with 95% confidence intervals. Proportions/frequencies, κ, and suitable simple effect sizes were among the evaluation metrics.

Validation & quality

High risk of bias studies was identified and tested using sensitivity analysis (re-synthesizing without those studies); quality appraisal used CASP/MMAT and was noted in the extraction sheet. The statistical presumptions (multicollinearity, fundamental goodness-of-fit) for association tests were examined. Listwise deletion was performed for descriptive variable missing data if the percentage was less than 5%; simple imputation (mode/majority category) was used for reporting frequencies but not for association tests if the percentage was greater than 5%. A qualitative analysis was conducted on outlier textual cases, such as unusual domains. A strategy codebook with annotated examples, PRISMA screening logs, analysis scripts with fixed random seeds, and archived search strings and search dates all assist reproducibility. Data/code availability: in accordance with institutional access requirements [repository link/access conditions], the Python scripts, PRISMA diagram, codebook, Excel corpus, and screening logs will all be placed in an open/restricted repository.

Human subjects ethics

Since all of the data used in this review came from publicly accessible publications, there is no direct use of human subjects; hence, the following participant consent statement is not applicable: “Before beginning the study, all subjects provided their informed consent for inclusion”. In order to comply, we either ask the institutional ethics committee for a Non-Human Subjects Research/exclude determination or rely on national legislation that exclude document based SLRs of published research (if the editor requests it, an exemption letter, number, and date will be provided). The consent declarations and anonymity guarantee provided by the original authors of any primary research in the corpus were recorded without any further communication. We took note of the safeguards mentioned in main studies that included vulnerable groups. On request, the editor can obtain all supporting documentation (protocol, search logs, screening forms, and codebook).

Finding and Discussion

A synthesis based on an SLR of 23 articles in an Excel dataset confirms that translation practices in EFL/ESL contexts are moving toward a hybrid ecosystem that combines classical strategies, the use of MT/GenAI, and *post editing* as the fulcrum of quality control. “English teaching focuses on three aspects. In the following, the focus of the study is the literary works that have been translated into the English language.” In line with the framework in the *Introduction and Method*, the synthesis shows a consistent pattern: MT/GenAI tools are positioned as draft accelerators, while the final quality especially in genres sensitive to style and culture such as literature is determined by strategically informed human

intervention at the *post editing* stage. This practice is reflected in classroom reports and applied studies that normalize the integration of Google Translate/DeepL and GAI tools into translator training, while still emphasizing human curation and evidence of outcomes at the unit of data level (Abdelhalim, Alsahil, & Alsuhaibani, 2025; Sinambela & Siregar, 2023; Varela Salinas & Burbat, 2023). At the same time, competence and culture oriented studies show that literary translation becomes an effective vehicle for cross-cultural mediation *provided that* micro strategies such as the explicitation of realia, compensation for imagery/rhythm, and idiomatic adjustment are operationalized explicitly within evaluation rubrics (Junaeny & Nirdayanti, 2023; Muñoz-Basols, 2023).

Furthermore, methodological patterns across the 23 articles reinforce this rationale: classroom research on CAT and MT applications tends to report efficiency gains and increased student confidence, while also acknowledging "surface fluency" that must be re-examined to prevent semantic ambiguity and to preserve the rhetorical purpose of the target text (Dorst et al., 2022; Liu, 2022; Yang, 2022). On the measurement side, the development of instruments and evaluative approaches ranging from *measurement tools* for dual-representation translation skills to questionnaire based assessments provides an objective footing to link outcome indicators (readability, equivalence, stylistic resonance, preservation of realia) with strategic steps at concrete segments (Erdogan, Çetin, & Ari, 2021; Kilaton, 2024). Accordingly, the study's second focus (literature → English) can be situated within a replicable instructional pathway: drafting via MT/GenAI → structured *post editing* based on a strategy taxonomy → *before–after* verification on sample data units → pedagogical reflection that consolidates novice translators' competencies a blueprint consistent with findings in business/academic contexts and with proposals to use AI for formative feedback without relinquishing the authority of human evaluators (Alkhbeer, 2025; Maid & Azmi, 2025; Wang, 2021).

Analysis of literary works translated into English

Within the corpus, the subset that explicitly mentions literature and English is indeed limited, yet its reporting pattern is consistent: study abstracts more often foreground outcome achievements readability, preservation of cultural nuance, and stylistic consistency rather than naming micro strategies at the level of data units. This "target language oriented" pattern can be understood as a consequence of the nature of literary texts, which prioritizes aesthetic resonance and semantic integrity as the main benchmarks, while also reflecting the integration of MT/GenAI tools that produce "surface fluent" drafts but still require human curation to remove ambiguities, reorganize information structure, and restore the author's idiolect (Abdelhalim, 2025; Varela Salinas & Burbat, 2023). In other words, when abstracts highlight outcome indicators, they are revealing the intersection between technological efficiency and translator craftsmanship: machines accelerate draft production, while post editing becomes the operational arena for strategy explicitation of cultural realia, compensation for imagery/rhythm, and idiomatic adjustment that ensures message, style, and rhetorical purpose remain intact in the target language (Alwazna, 2023; Junaeny & Nirdayanti, 2023).

This tendency also aligns with findings from classrooms and translator training programs, where the use of Google Translate/DeepL and other GAI tools

is “permitted” as an accelerator of idea generation, but the final quality is only assured when students carry out targeted post editing using rubrics that link outcome indicators to clear micro actions at the unit of data level (for example, before after exemplars for cultural allusions or conventional metaphors) (Muñoz-Basols, 2023; Varela Salinas & Burbat, 2023). In practice, this often appears as a shift in focus from merely “naming strategies” to “demonstrating the effects of strategies” on the output: rather than writing “explicitation” in the abstract, researchers show that target readers judge the post-edited version to be clearer, more “literary-sounding,” and more culturally familiar. In teaching contexts, this pattern has proven effective in fostering students’ strategic awareness they not only know what needs to be improved, but also how to improve it at the vulnerable points (Maid & Azmi, 2025; Sinambela, 2023).

The implication is that reporting standards for literature→English studies should bind outcome indicators to operational evidence at the segment level: source–target excerpts, the rationale for the chosen step (e.g., explicitation of a particular cultural referent), and its impact on readability, rhythm, and the narrator’s persona. Such reporting makes claims like “cultural nuance preserved” auditable across studies and facilitates replication in other genres without abandoning the efficiency brought by MT/GenAI (Abdelhalim, 2025; Omolu & Mappewali, 2024). Thus, even though strategy labels are often absent from abstracts, the corpus still supports a “draft by machine, craft by human” model that places strategically informed post editing at the core of quality for literary translation into English.

Table 1. Strategy in literary subsets → English

Strategy	Frequency in Abstracts
Literal	0
Adaptation	0
Domestication	0
Foreignization	0
Explicitation	0
Implicitation	0
Compensation	0
Equivalence	0

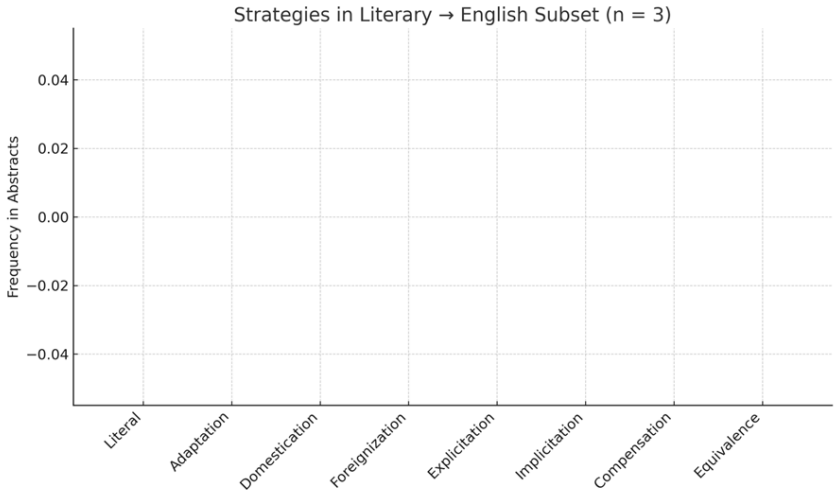

Figure 1. Strategies in a Subset of Literary Works → English

Table 1 and Figure 1. Strategies in the English Language Literary Subset summarize a scan of the strategy taxonomy (literal, adaptation, domestication/foreignization, explicitation/implicitation, compensation, equivalence) in the abstracts of the three subset articles: explicit strategy labels do not appear at the abstract level. This "target language" pattern is consistent with observational findings in translator training that "breaks the taboo" on using MT in the classroom, while still positioning post editing as the mechanism for quality control (Omolu & Mappewali, 2024; Varela Salinas & Burbat, 2023).

In the literary domain, *post editing* functions as an operational space for concrete and measurable strategies: the translator performs explicitation of cultural references (for instance, adding clarifying information to local realia or allusions) to prevent misinterpretation, applies compensation for imagery and narrative rhythm (replacing stylistic devices elsewhere when the original effect cannot be maintained literally) to preserve aesthetic resonance, and makes idiomatic adjustments so that the pragmatic effect in the target language matches that of the source. This mechanism operates alongside restructuring of informational flow and clarification of references in sentences or paragraphs that appear fluent on the surface but contain ambiguity; as a result, the global coherence and the author's "voice" are more consistently perceived by the target readers. Correspondingly, studies examining the outputs of ChatGPT/Google Translate on literary texts show a rise in acceptability after post editing even when initial lexico grammatical accuracy is relatively high because it is this post editing stage that restores the stylistic effects and cultural content not captured by mere formal correspondence (Abdelhalim, 2025; Alwazna, 2023). In pedagogical practice, this improvement becomes evident when evaluations are conducted using before after examples at the data unit level (sentence lines, dialogues, or key imagery) with rubrics combining readability, semantic equivalence, and aesthetic indicators; thus, *post-editing* is not merely linguistic correction but a form of literary effect engineering that bridges the efficiency of MT/GenAI and the quality demands of the genre.

Cross article findings also show that preserving cultural nuance matters more for literary texts than any single technical metric. Accordingly, pedagogical and applied studies emphasize exercises in analyzing cultural realia combined with targeted post editing strategies to maintain the author's "voice" and the genre's register (Junaeny & Nirdayanti, 2023; Maid & Azmi, 2025). In practice, students flag high risk segments (idioms, allusions, honorifics), choose appropriate treatments (e.g., explicitation + gloss, domestication/foreignization, stylistic compensation), and then verify the effect through before after examples at the unit-of-data level. In this way, post-editing functions as the engineering of cultural aesthetic effects that bridges the efficiency of MT/GenAI and the quality demands of literature.

On the other hand, the contribution of translation as a bridge for cultural learning strengthens the argument that literature→English requires a curriculum that integrates tool use (MT/GenAI for *drafting*) with human micro strategies at the *post-editing* stage. When translation is positioned as a practice of cultural mediation, students not only master lexical equivalents, but also develop sensitivity to realia, idioms, and genre conventions, especially if instruction is complemented by structured reflective tasks for example, *think alouds*, *translation commentary*, and reflective journals that require justifying strategic decisions at the unit of data

level. This approach becomes even stronger when paired with an annotated mini corpus (before after examples, cultural glosses) and rubric based peer feedback, enabling students to see directly how strategic choices impact readability and target reader reception (Abadou, 2024; Muñoz-Basols, 2023). Thus, the curriculum functions as an integrative ecosystem: the machine accelerates draft production, while strategy aware *post editing* documented in a portfolio consolidates the cultural mediation competence at the core of literary translation.

The distinctive features of literature such as an author's idiolect, intentional ambiguity, and prosodic rhythm are often "smoothed out" by MT drafts because models prioritize surface coherence and the conventions of general corpora. Hence, the pattern "draft by machine, craft by human" emerges strongly across corpora: the machine accelerates the preparation of an initial draft, while the human restores layers of meaning, style, and aesthetics through strategy aware *post-editing*. This restoration includes safeguarding the narrative voice (lexical choices consistent with the persona), mapping metaphors (replacing untransferable figurative language with stylistic compensation elsewhere), aligning rhythm (recalibrating clause length/enjambment for poetic prose/poetry), and reinforcing pragmatics (adjusting honorifics, *deixis*, and cultural implicatures). In practice, this is typically carried out through before after cycles at the unit-of-data level and supported by rubrics that assess readability, semantic equivalence, and stylistic resonance, so that the final result "sounds literary" to target readers consistent with observational and comparative findings that normalize the use of MT/GAI in the classroom while asserting the priority of human *post editing* to preserve literary effects (Abdelhalim, 2025; Varela Salinas & Burbat, 2023).

Classroom findings that employ applied translation likewise show gains in strategic competence when students work with rubrics that link outcome indicators to micro actions for example, before after comparisons on cultural allusions accompanied by justification of the chosen approach (explicitation/gloss, domestication foreignization, or stylistic compensation). Such gains typically arise within cycles of authentic tasks that include think-alouds, calibrated peer review (with anchor examples), and a revision portfolio, so that strategic transfer is tested across genres and levels of textual complexity. This approach aligns with the principles of situated learning competence is built through meaningful participation in the practices of the translation communityan d has been shown to strengthen students' procedural understanding as well as their cultural sensitivity (Yıldız, 2020). In EFL contexts, integrating rubrics with authentic tasks also correlates with increases in functional accuracy and linguistic confidence in real classroom situations, underscoring that instructional designs which tie micro-strategies to measurable outputs are effective for developing novice translators' competence (Sinambela, 2023).

Finally, the literary subset in the corpus implies reporting standards that need to be clarified: every claim that "cultural nuance is preserved" should be linked to concrete data units (sentences/paragraphs) along with naming the strategic steps taken for example, explicitation of realia, brief glosses, or stylistic compensation as well as the rationale for the choice and the intended effect on target readers. Ideal reporting includes a mini-format: *SL → TL (MT/GenAI draft) → TL (post–post-editing)*, complete with notes on strategies and the outcome indicators assessed (readability, cultural coherence, rhythm/style). To enhance auditability and cross-

study comparability, these items can be supplemented with strategy tags (e.g., EXP, DOM/FOR, COMP), inter-rater agreement tests, and contextual metadata (genre, language pair, learner profile). Such practices are especially relevant when MT/GenAI is part of the workflow, because they allow readers to evaluate the specific contribution of post editing rather than relying on general impressions (Omolu & Mappewali, 2024). In the classroom, the same approach namely, including strategy labeled before after examples on high-risk segments (idioms, allusions, cultural lexicon) has been shown to help students connect micro-actions to macro impacts on literary reception, strengthening the learning of reproducible strategies (Junaeny & Nirdayanti, 2023).

Analysis methodology and pedagogical impact

Table 2. Distribution of methods/designs in the corpus

Method Desind	Count
Survey Questionaire	13
Systematic Review/Review	18
Corpus/Comparative	7
Experiment/Quasi-Experiment	5
Mixed Methods	3
Classroom Study	2

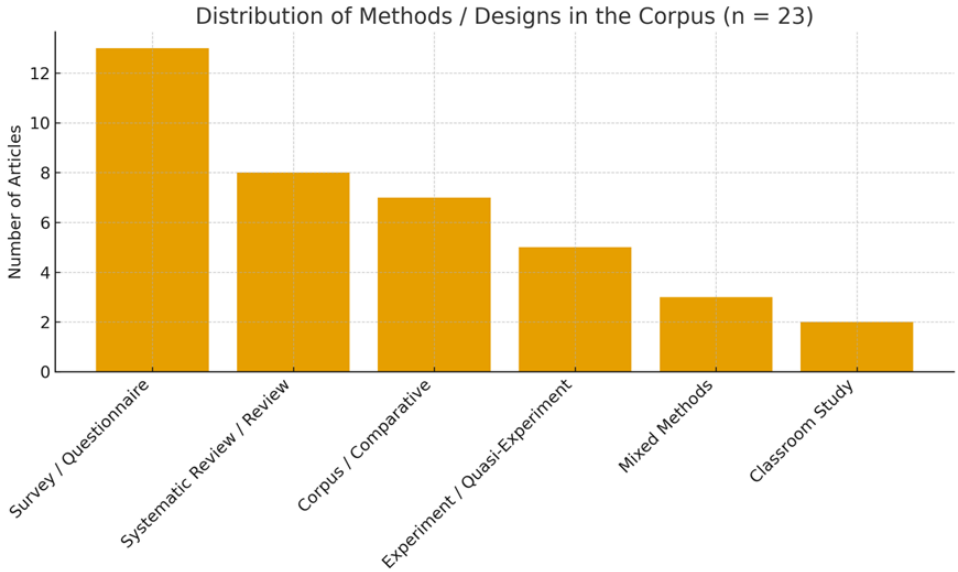

Figure 2. Distribution of methods/designs in the corpus

The methodological composition of the 23 articles shows a predominance of survey/observational and synthesis studies, followed by corpus/comparative studies and experiments/quasi experiments. This distribution explains why summaries often employ “outcome language” (readability, acceptability, nuance) rather than “strategy language” (naming micro-steps), while at the same time calling for a more explicit causal design to test the mechanisms of post-editing strategies (Liu, 2022; Yang, 2022).

Observational evidence shows that the adoption of MT/GenAI in classrooms is becoming increasingly widespread and, at the initial stage, tends to focus on perceived usefulness (comfort, speed, willingness to experiment) rather than on in-depth evaluation of output quality. In many contexts, Google Translate and GAI tools do indeed save drafting time and assist with initial idea generation, yet instructors consistently emphasize the post editing stage to maintain rhetorical purpose and clarity of meaning including restructuring, reference clarification, and register adjustment to fit genre and audience (Bin Dahmash, 2020; Dorst et al., 2022). To balance efficiency with quality, an effective practice is gradual scaffolding (draft → guided *post-editing* → *before–after* analysis at the data-unit level) accompanied by transparent rubrics and calibrated peer review, encouraging students to move beyond "surface fluency" toward functional and rhetorical equivalence in the target text (Liu, 2022; Yang, 2022). Within this framework, MT/GenAI serves as a process accelerator, while *post-editing* acts as a quality guarantor, linking micro level decisions to measurable learning outcomes.

In line with this, training based on computational resources including *coding* to "unpack" how MT works has been shown to enhance translation students' technological literacy while fostering a critical attitude toward machine output. Effective practices include using simple *notebooks* to run pre post processing pipelines (tokenization, *subwording*, rough *alignment*), conducting targeted error analyses (e.g., detecting referential ambiguity or cultural *named entities*), and implementing mini evaluation protocols (*before after* examples accompanied by strategic justification rather than mere automatic scores). The integration of practical repositories (e.g., small annotated datasets and evaluation scripts) and light modeling tasks such as *shadow MT* or controlled *prompt engineering* makes students better prepared to perform evidence-based *post editing*, as they can connect technical findings to micro strategic decisions at the data unit level (Krüger, 2021; Yang, 2022).

In the domain of general skills, translation is also used to integrate 21st-century skills collaboration, problem solving, and digital literacy through task designs that require team coordination, responsible use of MT/GenAI tools, and metacognitive reflection on strategic decisions. In practice, students work in different roles (translator, *post editor*, evaluator), combining technical findings with rhetorical considerations, and then write a *rationale* for their choices (e.g., explicitation/gloss or stylistic compensation) using before after examples. This pattern is perceived as an authentic medium for integrating language, technology, and self reflection enhancing confidence, process awareness, and accountability for outcomes as reported in project based and *applied translation* classroom contexts (Bilal, 2021; Muñoz-Basols, 2023).

Beyond the literary domain, the development of specialized assessment instruments also supports the objectification of quality for instance, measuring *multiple-representation translating skills* in the STEM domain by mapping transformation abilities (verbal ↔ symbolic ↔ graphical) to indicators aligned with *post-editing* rubrics (conceptual accuracy, representational cohesion, and clarity of reasoning). Such instruments allow for stronger content validity and higher inter-rater reliability because the criteria are made explicit at the data unit level (before after) and linked to typical MT error types symbolic *misalignment*, loss of inferential steps, or misinterpretation of graphs (Erdogan, 2021). When combined

with *post editing* tasks, cross representational indicators can be operationalized as formative checkpoints: students identify high-risk segments, carry out targeted corrections (glossing/explicitation of steps), and then verify improvements using the same rubric, making the feedback loop consistent and measurable (Kilaton, 2024). In this way, evaluation does not merely assess surface "right wrong" accuracy but measures the precision of meaning mapping across representations precisely the aspect that is often fragile in MT draft outputs.

AI based evaluation engineering has also emerged as an automated feedback mechanism that accelerates *closing the loop* between drafting, *post editing*, and formative assessment. The Transformer GAN approach combining contextual representation (Transformer) with *adversarial refinement* (GAN) provides scores and improvement suggestions aligned with quality indicators (readability, semantic equivalence, and stylistic coherence), which are then returned to students/instructors for human validation on critical data segments. This architecture is ideal for a hybrid model: AI performs *prescreening* of patterned errors (e.g., referential ambiguity, term *mistranslation*, or register deviation) and offers *micro edit* recommendations, while instructors maintain final judgment, assessment fairness, and accountability Alkhbeer (2025). On the adoption side, user attitude studies indicate that acceptance of MT/GAI increases when the tools are tied to transparent rubrics and a revision audit trail; in other words, AI output is not a replacement but an evaluation companion that remains compatible with manual *post editing* and curricular standards (Liu, 2022).

Evidence from Hong Kong and European contexts shows that instructors and learners attitudes toward MT are positively critical: the tools are used for efficiency (saving drafting time, exploring equivalents), but the final quality is still guaranteed through strategic literacy the ability to select, apply, and justify micro level steps during *post editing* (Dorst et al., 2022; Liu, 2022). These findings reinforce a policy of "responsible accompaniment" rather than prohibition: institutions encourage the use of MT/GAI within rubric-bound workflows (outcome indicators micro actions), revision audit trails (*pre-/post-editing* logs), tool disclosure in assignment sheets, and calibrated peer review to check register coherence and referential clarity. Within this framework, MT serves as a process accelerator, while strategic literacy ensures functional and rhetorical equivalence in the target text maintaining academic quality without closing off pedagogical innovation.

Pedagogically, the implications point to an operational and auditable replicable instructional cycle:

(1) Drafting via MT/GenAI to reduce initial workload and generate candidate equivalents;
(2) Structured post-editing based on a strategy taxonomy (explicitation/gloss, domestication foreignization, stylistic compensation, idiomatic adjustment) carried out in assigned roles such as translator, *post editor*, and evaluator;
(3) Checking for structural ambiguity and referential clarity using checklists (subject predicate relations, pronominal reference, *deixis*, register) and before after comparisons at the data unit level to assess the impact of micro-actions; and

(4) Evaluative reflection guided by rubrics linking readability, semantic equivalence, stylistic resonance, and preservation of cultural realia with the strategic decisions made.

Practical implementation includes calibrated peer review, revision audit trails (logging tools and strategic steps), and short feedback loops that map findings to subsequent improvements. This sequence aligns with emphases in tool-adoption studies which position AI as an initial feedback provider while final judgment remains human as well as with applied evidence across contexts showing that CAT/MT integration is effective when tied to transparent rubrics and formative evaluation based on data segments (Alkhbeer, 2025; Yang, 2022).

To make outcome claims auditable, each rubric indicator must be directly linked to micro actions at the data-unit level for example, the handling of cultural transfer, metaphor, or register and supported with before after examples that include both the rationale for the chosen strategy and the intended effect on the target reader. To strengthen traceability, examples can be assigned strategy tags (e.g., EXP for explicitation, DOM/FOR for domestication–foreignization, COMP for compensation) and recorded in an audit sheet that maps "indicator → text segment → action → change." Assessment is then complemented with calibrated peer review and, when possible, brief inter-rater reliability checks on a subset of segments. Through this workflow, the relationship between strategy and outcome becomes visible and measurable, while also facilitating cross class/genre replication, since each improvement can be emulated in similar cases (Junaeny & Nirdayanti, 2023; Yildiz, 2023).

At the course level, fostering responsible prompting for instance, specifying the translation's purpose, audience, register, and cultural constraints within the *prompt* alongside logging of the post editing process (what the tool generated, which micro-strategies were applied, and the rationale for revisions) helps instructors monitor progress while maintaining academic integrity. This practice is typically paired with tool-use declarations on assignment sheets, revision audit trails (before after traces labeled with strategies), and transparent rubrics that link outcome indicators to micro actions; this combination shifts the focus from merely "using the tool" to accounting for strategic decisions. Consistent with observational evidence, a "guided support" approach rather than prohibition is more effective in directing MT/GenAI toward productive learning pathways: the tools are used to accelerate drafting and explore equivalents, while the final quality is ensured through human *post editing* that is strategy aware and auditable (Bin Dahmash, 2020; Varela Salinas & Burbat, 2023).

When dealing with business and non-literary texts, a focus on procedural competence terminology management, lexical cohesion/collocation consistency, formatting compliance, and adherence to a house *style guide* remains crucial; however, the "lessons from literature" regarding sensitivity to nuance have proven to enhance final quality, particularly in maintaining professional *tone*, rhetorical clarity, and precision in cultural implicatures. In practical terms, training that explicitly teaches strategies (explicitation/glossing of culturally loaded terms, register alignment, stylistic compensation when literal equivalence reduces persuasiveness) and demonstrates them through before after examples improves both readability and the credibility of business discourse (Wang, 2021). Findings

from the French English language bridge likewise show that modules combining a verified terminology bank, genre templates (emails, reports, proposals), and guided post-editing sessions foster stable strategic transfer enabling students to choose equivalents more accurately, maintain corporate tone, and avoid cross linguistic interference bias (Maid & Azm, 2025). Thus, the integration of procedural discipline with sensitivity to nuance forms the core of quality in effective and persuasive business/nonliterary texts.

Overall, the evidence map affirms the "draft by machine, craft by human" model, supported by two pillars: technological literacy (understanding the capabilities and limitations of MT/GAI, responsible prompting, and revision audit trails) and strategic literacy (the ability to select, apply, and justify micro-level actions during *post-editing*). In the literary domain, this combination preserves the idiolect, rhythm, and cultural realia through *before after* procedures that link outcome indicators to micro-actions, ensuring that the output still "sounds literary" to the target audience (Sinambela, 2023). In the non-literary domain, the same framework enhances procedural reliability maintaining consistent terminology, lexical cohesion, and genre appropriate formatting while achieving efficiency without sacrificing semantic accuracy; here, *post editing* focuses on rhetorical clarity and functional equivalence of message (Guo, 2023). Thus, MT/GAI is positioned as a draft production accelerator, while human *post-editing* serves as the quality guarantor bridging technological speed and academic/professional standards through transparent, auditable evaluation rubrics applicable across contexts.

As a closing synthesis linking back to the research aims, these findings suggest three mutually reinforcing operational steps. First, standardize *before after* reporting and the naming of micro-strategies at the data unit level (e.g., tags EXP/DOM FOR/COMP) along with the rationale for selection and the intended effect, so that every outcome claim can be audited and compared across studies or classes. Second, increase the number of causal designs (experiments/quasi-experiments with *pre post* and inter-rater evaluation) and corpus-based studies that connect micro level changes lexical, syntactic, pragmatic to outcome indicators (readability, semantic equivalence, stylistic resonance, preservation of realia), allowing recommendations to move from descriptive to evidence-based prescriptive. Third, institutionalize a replicable instructional cycle that integrates MT/GenAI as a draft accelerator with guided *post editing* based on transparent rubrics, revision *audit trails*, and calibrated *peer review*. This framework aligns with the mandate to strengthen novice translators' competence enhancing both technological and strategic literacy while meeting the need for an evidence based curriculum that leverages computational resources without relinquishing human judgment (Alkhbeer, 2025; Krüger, 2021).

Conclusion

This systematic literature review demonstrates that translation practices in EFL/ESL contexts are evolving toward a hybrid model that balances technological efficiency with human creativity and cultural sensitivity. The synthesis across 23 studies confirms the centrality of post-editing as a strategic and pedagogical bridge between MT/GenAI outputs and high-quality translations that preserve meaning, style, and cultural nuance. Evidence from literary, business, and classroom-based

research consistently supports the "draft by machine, craft by human" paradigm, where MT/GenAI accelerates the drafting process while human translators ensure rhetorical, aesthetic, and cultural integrity through informed post editing. Pedagogically, this integration fosters both technological and strategic literacy, empowering learners to understand not only how to use translation tools effectively but also why specific micro strategies such as explicitation, compensation, and stylistic adjustment matter for quality outcomes. The findings advocate for standardized before after reporting formats, transparent rubrics linking micro-actions to outcome indicators, and replicable instructional cycles that combine efficiency, accountability, and critical reflection. Ultimately, the review highlights that responsible synergy between human and machine translation practices can enhance translation education and professional standards while preserving the interpretive and ethical dimensions that define the human craft of translation.

References

Abadou, F. (2024). Investigating the role of translation in teaching culture to foreign language learners at institutions of higher education in Algeria. *Turkish Academic Research Review*, *9*(3), 259-276. https://doi.org/10.30622/tarr.1497376

Abdelhalim, S. M., Alsahil, A. A., & Alsuhaibani, Z. A. (2025). Artificial intelligence tools and literary translation: A comparative investigation of ChatGPT and Google Translate from novice and advanced EFL student translators' perspectives. *Cogent Arts & Humanities*, *12*(1), 2508031. https://doi.org/10.1080/23311983.2025.2508031

Akhan, O., & Uzun, A. (2025). Developing historical thinking skills and creativity of visually impaired middle school students. *Frontiers in Psychology*, *16*. https://doi.org/10.3389/fpsyg.2025.1509297

Alkhbeer, A. T. (2025). A transformer-GAN based approach for evaluating and improving high school students' English translation skills. *Journal of Information Systems Engineering and Management*, *10*(36), 938–955. https://doi.org/10.52783/jisem.v10i36s.6615

Alwazna, R. (2023). The use of translation theory through reconciling between Englishisation and translanguaging by Arab instructors in EMI higher education classes: Training postgraduate students to be translators …. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.1010704

A.R., R., & Pawar, R. (2020). Exploring the role of artificial intelligence (AI) on student learning. In H.K. Yadav & P.K. Yadav (Eds.), *Artificial intelligence and ChatGPT (A transformative approach)*. https://doi.org/10.25215/9392917538.21

Bilal, D. M. (2021). International students' acquisition of library research skills: relationship with their English language proficiency. In M. Pastine & B. Katz (Eds.), *Integrating Library Use Skills into the General Education Curriculum* (pp.129–145). New York: Routledge. https://doi.org/10.4324/9781315859835-15

Bin Dahmash, N. (2020). I can't live without Google Translate: A close look at the use of Google Translate app by second language learners in Saudi Arabia. *Arab World English Journal*, *11*(3), 226–240. https://doi.org/10.24093/awej/vol11no3.14

Dorst, A. G., Valdez, S., & Bouman, H. (2022). Machine translation in the multilingual classroom: How, when and why do humanities students at a Dutch university use machine translation? *Translation and Translanguaging in Multilingual Contexts*, *8*(1), 49–66. https://doi.org/10.1075/ttmc.00080.dor

Erdogan, S. M., Çetin, H., & Ari, K. (2021). Development of multiple representation translating measurement tool and examination of 9th grade students' multiple representation translate skills in Algebra. *Acta Didactica Napocensia*, *14*(2), 160-180.

Guo, W. (2023). Analysis of the key points of improving students' English application ability in English translation teaching. *Contemporary Education and Teaching Research*, *4*(1), 1–4. https://doi.org/10.47852/bonviewcetr2023040101

Junaeny, A., & Nirdayanti, N. (2023). Analysis of students Chinese-Indonesian translation of cultural terms translation. *Sinolingua: Journal of Chinese Studies*, *1*(2), 91–91. https://doi.org/10.20961/sinolingua.v1i2.71964s

Kilaton, L. (2024). Assessing proficiency level of research questionnaire translation skills in higher education students. *Psychology and Education: A Multidisciplinary Journal, 20*(3), 297-306. https://doi.org/10.2139/ssrn.4845063

Krüger, R. (2021). An online repository of python resources for teaching machine translation to translation students. *Current Trends in Translation Teaching and Learning E*, *8*, 4–30. https://doi.org/10.51287/cttle20212

Liu, K., Kwok, H., Liu, J., & Cheung, A. (2022). Sustainability and influence of machine translation: Perceptions and attitudes of translation instructors and learners in Hong Kong. *Sustainability, 14*(11), 6399. https://doi.org/10.3390/su14116399

Maid, Y., & Azmi, N. (2025). Bridging the linguistic gap: Translation skills for Moroccan business students in the transition from French to English. *European Modern Studies Journal*, *9*(3), 265–275. https://doi.org/10.59573/emsj.9(3).2025.22

Muñoz-Basols, J. (2023). Potentialities of applied translation for language learning in the era of artificial intelligence. *Hispania*, *106*(2), 171–194. https://doi.org/10.1353/hpn.2023.a899427

Omolu, F. A., & Mappewali, A. (2024). The impact of translation tools towards student translators' skills. *Premise: Journal of English Education*, *13*(1), 340–340. https://doi.org/10.24127/pj.v13i1.8731

Sinambela, E., Siregar, R., & Pakpahan, C. (2023). Improving students' ability in using English with a simple translation: A case on elementary school level. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, *7*(3), 3267–3278. https://doi.org/10.31004/obsesi.v7i3.4647

Varela Salinas, M.-J., & Burbat, R. (2023). Google Translate and DeepL: Breaking taboos in translator training: Observational study and analysis. *Ibérica*, (45), 243–266. https://doi.org/10.17398/2340-2784.45.243

Wang, A. (2021). A study on the problems and skills of business English translation. *Learning & Education*, *10*(3), 186. https://doi.org/10.18282/l-e.v10i3.2444

Yang, H., Aniceto, N., Allen, C., Bulusu, K. C., & Tsang, S. L. (2022). Predicting anticancer synergistic activity through machine learning and natural language processing. In H. Yang (Ed.), *Data science, AI, and machine learning in drug development* (pp. 95–120). New York: Chapman and Hall/CRC. https://doi.org/10.1201/9781003150886-5

Yildiz, T. (2023). Measurement of attitude in language learning with AI (MALL:AI). *Participatory Educational Research*, *10*(4), 111–126. https://doi.org/10.17275/per.23.62.10.4