

When Support Becomes Dependency: Investigating ChatGPT’s Role in Shaping EFL Students’ Translation Competence

Tran Thi Ngoc Lien*

Haiphong University of Management and Technology, Vietnam

Correspondence

Email: lienttn@hpu.edu.vn

Abstract

Integrating artificial intelligence into translation studies has proven to enhance learners’ translation skills and practice. However, overreliance on this technological tool may give rise to various translation-related issues. This study, therefore, investigates both the benefits and potential risks associated with the use of ChatGPT in translation training for Vietnamese English-major students. It also aims to propose some pedagogical approaches that position AI as more of a scaffolding tool than a substitute in translation training at a university in Vietnam. Sixty third-year English-major students participated in the study and were divided into two groups, with the control group completing the translation tasks manually while the experimental group utilized ChatGPT for both learning and assessment. Pre-tests and post-tests were used, and the assessment was based on EMT (European Master’s in Translation) dimensions, including language, intercultural, thematic, and strategic competences. Semi-structured interviews were also carried out to explore student perception about the intervention. Findings indicate that ChatGPT improved translation quality by enhancing students’ language, thematic, and strategic competencies; however, it did not really influence their intercultural competence. The problems of pseudo-efficiency and learner autonomy remain a concern.

ARTICLE HISTORY

Received: 15 August 2025

Revised: 20 November 2025

Accepted: 08 December 2025

KEYWORDS

ChatGPT, Translation
Competence, EMT
Framework

How to cite this article (APA 7th Edition):

Lien, T. T. N. (2025). When support becomes dependency: Investigating ChatGPT’s role in shaping EFL students’ translation competence. *Language Testing in Focus: An International Journal*, 12, 1–20. <https://doi.org/10.32038/ltf.2025.12.01>

Introduction

Artificial intelligence (AI) such as ChatGPT has exerted great impacts on translation practice and pedagogy, not to mention its capability to change the way translators work. Studies have confirmed that this chatbot is capable of rendering different types of texts with a high degree of accuracy (Alosaimi & Alawad, 2024; Chen, 2024; Chen

& Lin, 2025; Jiao et al., 2023; Nguyen, 2025; Nguyen, 2024; Sun, 2024; Zhao, 2025;). It is ChatGPT's ability to offer students with instant lexical assistance, stylistic refinement, and fluency enhancement that supports the translation learners to work faster and more effectively. However, this tool has also revealed certain downsides when it comes to the translation of culture-bound texts such as literary works, which "demands a high level of linguistic and cultural sensitivity" (Zhao, 2025, p.37), phrasal verbs whose translation requires a high level of clarity and accuracy as ChatGPT's translation finds it hard to "fit the context or paint a clear picture of the intended meaning to the audience" (Alosaimi & Alawad, 2024, p.289), or engineering texts which often requires "amendments in terminology translation, format adjustment, and language localization" (Zhou, 2025, p.61).

Although ChatGPT has long been considered as an effective pedagogical tool, its application has recently become prevalent. The studies of its overall effects on improving students' EFL proficiency in general and their translation competence in particular have been far from sufficient in Vietnam. As translation courses in the Vietnamese EFL training context are predominantly traditional, relying on manual translation, dictionary use, and teacher correction, there is a need to take the role of AI incorporation into more serious consideration, especially when AI tools have quickly expanded beyond their embryonic stage of integration into translation training programs. This study, therefore, aims to investigate both the benefits and risks of integrating ChatGPT into translation training courses for Vietnamese undergraduates at HPU, a university in a northern city of Vietnam. Drawing on the EMT 2022 framework, it aims to determine whether ChatGPT can be incorporated officially into translation classes more as a scaffolding tool than a substitute.

Literature Review

Translation Competence and its Conceptual Models

Translation competence (TC) has been recognized as a multifaceted construct. As early as the 1970s, Nida (1976) suggested that translators must know well the "subject matter" and "have some facility of expression in the receptor language" (p.47). Evaluating a professional translator, Wilss (1976), moreover, emphasized the mastery of both "linguistic and psycholinguistic" properties (p.119). Bachman (1990) as cited in Hatim & Mason (1997, p.200) stressed that "students are evaluated in terms of their relative degree of mastery of course content" and Bell (1991) recognized a translator as a communicator, emphasizing that translator competence is assessed through "five different kinds of knowledge; target language knowledge, text-type knowledge, source language knowledge, subject area ('real-world') knowledge; and contrastive knowledge" (p.36). Reiß (1971), cited in Hewson (2011), recommended a three-dimensional approach, which combines the analyses of text-type, linguistic components, and extra-linguistic determinants (p.3), and Hewson himself, when modelling translation criticism, referred to both linguistic and cultural knowledge of a translator (Hewson, 1995). As linguistic competence was no longer the only factor decisive for the assessment of a translator's ability, TC meant more than the

competence to work with language itself but to deal with the situational determinants. Various as it is, the definition of translator or translation competence highlights the ability to process a source language and translate it into a target language, concerning the transference of linguistic and extra-linguistic elements from one language into another (Newmark, 1988; Bell, 1991; Hewson, 1995; Hatim and Mason, 1997; Pym, 2003). The models for translation competence assessment, accordingly, have been developed to meet such requirements.

Among the most popular frameworks for assessing translation competence, PACTE (2017) gained worldwide recognition. It evaluated TC through five interdependent abilities: (1) *Bilingual competence*, indicating the control of pragmatic, sociolinguistic, and textual conventions in both source language and target language; (2) *Extra-linguistic competence*, reflecting the cultural knowledge that is required; (3) *Instrumental competence*, or the ability to use documentation sources, terminology databases, and technological tools; (4) *Knowledge-about-translation competence*, representing translators' awareness of theories, norms, and ethics; and (5) *Strategic competence*, showing the metacognitive ability to plan, monitor, and evaluate translation decisions. Despite its working strengths, this model has also faced critical scrutiny.

Other models are known for their outstanding defining qualities. Pym (2003) assessed TC as the one with "associative competence" and "macro-strategy," which exhibits the translators' mental competence and real-life experience in the translation process. Göpferich's (2009) *TransComp* model conceptualizes TC as a developmental system that evolves through iterative experience, reflection, and feedback. Kiraly (2014) proposed a framework grounded in a socio-constructivist standpoint emphasizing collaborative learning and socio-cognitive ability, whereas Schäffner & Adab (2000) highlighted the interplay between linguistic transfer and communicative purpose.

These theoretical perspectives converge on the fact that translation competence evaluation looks into not only pure linguistic proficiency but also a metacognitive capacity.

Translation Competence Assessment and EMT Framework

Translation competence assessment has gradually evolved from a product-oriented process to a multidimensional, process-based judgment. Early models, such as the one proposed by Kelly (2005), assess translator performance across cognitive, procedural, and attitudinal dimensions, or the Cognitive Diagnostic Assessment model by Leighton & Gierl (2007) was "designed to measure specific knowledge structures and processing skills in students to provide information about their cognitive strengths and weaknesses" (p. 3). Process-based frameworks (Angelone, 2010; Göpferich, 2009) were designed with keystroke logging, eye-tracking, and think-aloud techniques to handle translators' problem-solving strategies in real-life situations.

As a holistic model for translation assessment, the EMT framework was designed in 2009 with six defining domains, say, language competence, intercultural competence, information mining competence, thematic competence, and technological competence, working together to fulfil the task of translation service provision, the core of translation training objective (Gambier, 2009). The framework was revised in 2017, and the number of domains was reduced from "six" to "five". A total of 35 subskills are included in the updated set. It was renewed once more in 2022 with some changes and the addition of one new sub-competence.

Despite some improvements and alterations in terminology, EMT, like other translation assessment models such as PATCE (2017), puts not only language assessment but also other needs for translation training into perspective. Notably, the thematic competence delves into the translator's ability to update their knowledge about a specialist field, empowering learners' compatibility with the labour market (Esfandiari et al., 2019; Schäffner, 2012) or, it, in other words, can work well in an "academic environment with a wealth of professionally-oriented practices [...] to bridge the gap for students between academic and professional life" (Rothwell & Svoboda, 2019, p.53). The framework is also highly appreciated for its focus on "language services in a professional context" (EMT, 2022). It is, without doubt, that attention to translators' ability to integrate thematic expertise with language, intercultural, and technological competences is more likely to affect curriculum design and facilitate academic professional alignment (Jalambo et al., 2023). Another feature is its recognition of strategic competence, which requires learners to complete a process of translating, checking, reviewing, or revising, thus fostering learner autonomy (EMT, 2022; Plaza-Lara, 2022).

Despite its detailed evaluation of translation competence, the EMT framework has been accepted with critique. Completing 35 or 36 subskills must be an uphill task for translation learners, particularly those with low English proficiency. Regarding cultural and intercultural dimensions, the framework remains Eurocentric, emphasizing European multilingualism rather than global translation ecologies. Moreover, the assessment is completed via a set of multi-competencies supported with a more specific rubric. Normally, one sub-competence or skill presented in the framework comprises more than one action to be taken, which poses a great challenge for the evaluation work.

In this study, the EMT framework is applied with a simplified and adapted version. Five dimensions, which are language, intercultural, technological, thematic, and strategic competence, are evaluated based on 25 sub-competences.

Table 1*Translation Competence Assessment Dimensions (Adapted from EMT 2022)*

Analytical Dimension		Sub-competences <i>Students have the ability to:</i>
Language Competence	Sub1	produce grammatically accurate target texts;
	Sub2	use proper lexical equivalents in translation;
	Sub3	produce stylistically and semantically appropriate target texts;
	Sub4	use proper cohesive devices to reproduce the target texts;
	Sub5	generate textual equivalence;
Intercultural Competence	Sub6	be aware of cultural nuances;
	Sub7	realize the connotational meanings;
	Sub8	render pragmatic equivalence;
	Sub9	gasp contextual meanings;
	Sub10	exhibit the adaptation of tone, register, and communicative intent;
Thematic Competence	Sub11	develop understanding of subject matter and accurate transfer of conceptual meaning;
	Sub12	develop and use thematic and domain specific knowledge relevant to translation needs such as terminology, phraseology, specialized sources;
	Sub13	realize the specific situations, recipients, and constraints;
	Sub14	produce a 'fit for purpose' translation;
	Sub15	use text typology effectively;
Technological Competence	Sub16	use the most relevant IT applications, including the full range of office software, and adapt rapidly to new tools and IT resources having critically assessed their relevance and the impact of change on their work practices
	Sub17	Make effective use of search engines, corpus-based tools, text analysis tools, computer-assisted translation and quality assurance tools where appropriate
	Sub18	use ChatGPT ethically and efficiently to enhance translation productivity
	Sub19	justify the use of technological tools to support translation process
	Sub20	apply other tools in support of language and translation technology, such as workflow management tools
Strategic Competence	Sub21	check, review, revise and evaluate their own work and that of others according to standard or work-specific quality objectives and assess the appropriateness of using tools for the work at hand
	Sub22	analyse and justify their translation solutions and choices, using the appropriate metalanguage and applying appropriate theoretical approaches
	Sub23	understand and implement quality control strategies, using appropriate tools and techniques
	Sub24	edit source text content for processing using a translation tool
	Sub25	make a comparison of the translation done by technical tools and by human.

*ChatGPT and Translation Training**The application of ChatGPT to translation learning and assessment*

The adoption of ChatGPT in translation pedagogy has generated both enthusiasm and caution among educators and researchers. As recent scholarship demonstrates, ChatGPT's strengths lie primarily in its capacity to provide immediate feedback, lexical variety, and process efficiency (Dinh, 2025; Kim, 2018; Xiao & Zhi, 2024). Dinh's study found that ChatGPT could enhance students' critical engagement in their translation tasks and thereby improved their "translation accuracy, fluency, and vocabulary enrichment" (Dinh, 2025, p.114). Such findings resonated with Nguyen

(2025), who agreed that AI empowered learners to improve their language proficiency, translation skills, and other soft skills, such as teamwork or research abilities (p.15). Similarly, Nguyen (2024, p. 604) stated that ChatGPT facilitated text analysis, helped maintain terminology consistency, and offered immediate error-based feedback. The power of AI-assisted tools was also confirmed by Yusuf and Sivanadhan (2025). These scholars claimed that GPT-based systems revolutionized neural machine translation as they managed to integrate contextual awareness, improve fluency, and, to some extent, could work with idiomatic expressions more effectively than previous models.

Although ChatGPT has been hailed as one of the most widely used AI tools for EFL learning and translation training, its limitations still remains, ranging from its incapability to render cultural and contextual meanings to its failure to cultivate the actual translation competence for learners. According to Dinh (2025), ChatGPT expedited tasks but it could not convey connotational meanings or cultural references accurately, which, therefore, required extensive human correction. Similarly, Zhao (2025) noted that the model performed poorly on metaphorical or literary texts. Cognitively, continuous reliance on ChatGPT can lead to pseudo-efficiency, a great concern for translation pedagogists. Karataş et al. (2024) observed that while AI tools boosted motivation, they also risked fostering academic passivity. Learners though instructed to use AI as a supportive educational tool counted on it as soon as translation tasks are given, skipping analytical stages and setting aside the acquired skills and strategies for translation (Bui & Tran, 2024; Dinh, 2025).

ChatGPT Use through the Lens of EMT Competences

The integration of ChatGPT into translation training courses can be critically examined through four core dimensions of the EMT framework. They are linguistic, cultural, thematic, and strategic competences, each of which reflects both the potential and pitfalls of generative AI in translation.

First, using ChatGPT for translation tasks can enhance learners' language competence. This AI tool is confirmed to generate grammatically accurate English output. Deng (2024) concluded that ChatGPT translators outperformed human translators in language use as the number of errors made by the former was only half that of the latter (p.12). Echoing this, Dinh (2025) reported that Vietnamese students who used this app for their translation practice achieved higher surface accuracy and stylistic appropriateness. Zhao (2025) stated ChatGPT "excels in handling simple sentence structures" (p.36) while Haryanti (2024) hailed ChatGPT as a "state-of-the-art language model" which could "generate translations with a high degree of accuracy" (p.219). These findings are resonated in other empirical research, which confirmed the supportive function and capability of AI chatbots in (Alosaimi & Alawad, 2025; Khoshafah, 2021). However, ChatGPT's language competence has remained controversial. It shows certain limitations when working with "complex sentence structures, ambiguous words, and highly context-dependent texts" (Zhao, 2025, p. 36). Also concerning the downsides of using ChatGPT in translation, Zhou (2025)

posited that ChatGPT, though efficient in analyzing grammatical and semantic patterns much like a human, lagged behind in the preservation of informational essence and coherence. Fan et al. (2023) likewise raised a question about the “knowledge security risks” of using ChatGPT because its algorithm rules and pre-trained corpus did not show the superiority and subtlety to the perception of the complicated and abstract world (p. 50).

Second comes the matter of handling cultural sensitivity, or in other words, showing *cultural competence* by AI-driven systems like ChatGPT. Despite its ability to access extensive multilingual corpora, ChatGPT is widely acknowledged to struggle with culture-bound expressions (Banat & Adla, 2023; Chen & Lin, 2025; Dinh, 2025; Huang et al., 2020; Khoshafah, 2021; Macketanz et al., 2017; Sahari et al., 2023; Zhao, 2025). EFL learners relying on this tool may neglect questioning cultural appropriateness, leading to pragmatic flattening and loss of nuance.

Chen and Lin (2025) stated that AI has problems working with “specialized domains such as tourism, where texts often demand precise translations that reflect cultural nuances, idiomatic expressions, and regional language variations” (p.1). Similarly, Nguyen (2024) suggested that ChatGPT failed to convey all the intended nuances and “transfer metaphorical meanings” (p.608), or Chen & Lin (2025) admitted that ChatGPT “inadequately adapts culturally nuanced content” (p.3). Although ChatGPT is constructed with a huge pre-trained corpus and advanced algorithm, it cannot work effectively with cultural equivalents (Huang et al., 2020) or translate highly contextually dependent texts with “specific jargon and terms” (Khoshafah, 2021, p.3). As the translation has required more than just the competence of language use, tech-driven apps like ChatGPT face challenges handling the tasks that require aesthetic taste, cultural literacy, and social cognitivism (Fan et al., 2023).

Problematic as it is, the application of AI tools in translation has revealed a rosy outlook. Yusuf and Sivanadhan (2025) concluded from their findings that the improvement of AI translating idiomatic expressions and its ability to “maintain the stylistic nuances of the original texts” (p. 249). Echoing this viewpoint, Ghassemiazghandi (2024), who evaluated ChatGPT’s translation accuracy, revealed that ChatGPT-4’s translation managed to work with a literary text, a poem, and well preserve its format, and accurately render rhythms and rhymes, the core of a poem. This indicates that this chatbot has made progress in expressing “linguistic nuances, cultural context, and semantic meaning” (Ghassemiazghandi, 2024, p. 993). The futuristic potential of ChatGPT’s translation is also confirmed by Chen and Lin (2025), who worked out its better balancing of language accuracy and “cultural sensitivity and persuasiveness” (p. 13).

Another concern about the incorporation of AI tools, especially ChatGPT, into translation pedagogy is whether it can foster thematic competence. Within the EMT framework, the thematic competence refers to translators’ mastery of subject-specific

knowledge, terminology, and disciplinary conventions. Many scholars agreed that the application of translation technological models fostered students' thematic competence. Nguyen (2024) stated that ChatGPT could help students work with "specialized terms, while also reinforcing their topic-related knowledge" (p.608). This viewpoint, however, contrasts with Nguyen (2025), who raised concern about students having problems working with "terminology and vocabulary" as well as "understanding source materials" (p. 15).

The final aspect to be taken into consideration is the measurement of strategic competence. Put simply, translators who manifest a full capacity of this competence know how to self-control their work, revise and edit it, and create continuous improvement. This means that learners' autonomy can be enhanced as the use of ChatGPT encourages students to get engaged more in the translation process, supports them, and enables them to "learn from mistakes" (Haryanti, 2024, p.222), alongside promoting their confidence and self-regulation (Nguyen, 2024; Xiao & Zhi, 2024)

ChatGPT, supportive as it is claimed, may turn out to be counter-productive because learner autonomy might regress into detachment, where overreliance on automatic translation weakens intellectual engagement and problem-solving skills (Fan et al., 2023). As far as EMT's four competence dimensions are concerned, ChatGPT can function as a double-edged pedagogical innovation which helps translation trainees to better their linguistic, thematic, and strategic competences but not cultural sensitivity and autonomy.

Methods

This study adopted a mixed-methods quasi-experimental design to examine the impact of ChatGPT on EFL students' translation competence. It involved 60 third-year English major students from Haiphong University of Management and Technology. All had completed a basic translation course of 45 periods and exhibited intermediate-to-advanced English proficiency. They were divided into two groups of equal size. The experimental group (called GPTgroup in this study) was allowed to use ChatGPT for their translation tasks, whereas the control group (or MAN group) completed without technological support.

Quantitative analyses measured differences in translation performance between AI assisted students and those who did the translation manually. Qualitative analyses, on the other hand, explored students' perception and their possible dependency. The integration of numerical and narrative data supported a more holistic understanding of both product quality and the cognitive processes underlying translation (Creswell & Plano Clark, 2018).

Pre-tests and Post-tests were used to collect quantitative data, checking the first three competences of a translator, which are language competence, cultural competence, and thematic competence. All translation outputs were collected and coded as GPT1 to GPT30 (experimental) and MAN1 to MAN30 (control). The rating scale for each

sub-competence was from 0 to 5, and fractional scores are allowed in increments of 0.5. Following is the marking scale explanation of each sub-competence.

- 0: indicates no competence demonstrated;
- 1: indicates a limited control of the competences
- 2: indicates a control of the competences to some extent, but many language-in-use errors exist.
- 3: indicates a control of the competences, but some language-in-use errors exist. The ability to handle subtle nuances and context-specific knowledge is limited.
- 4: indicates a good control of the competences. Language-in-use errors are almost free; however, there still exist minor problems working with subtle nuances and showing context-specific knowledge.
- 5: indicates a good control of the competences and good ability to handle subtle nuances, and show context-specific knowledge. Language-in-use errors are almost free.

The quantitative data were analyzed using SPSS 26.0. A Paired Samples Test was used to interpret the differences in mean scores across groups. The researcher also employed semi-structured interviews to collect qualitative data. They were conducted with five randomly selected students who were coded IN1–IN5. The qualitative data were transcribed and thematically coded following Braun and Clarke’s (2019) six-phase framework. Codes were then mapped against EMT sub-dimensions to explain the mechanisms underlying quantitative findings. Triangulation ensured validity across both strands. The interview questions fell into six sub-categories including the four core dimensions in EMT, the warm-up section, and the reflective wrap-up (table 2).

The English–Vietnamese translation is a passage from the UNESCO Universal Declaration on Cultural Diversity (2001), while the Vietnamese–English translation is a short text on President Hồ Chí Minh’s views on culture. Both texts were approximately 250 words and comparable in lexical density and conceptual complexity. The experimental group was allowed to use ChatGPT (GPT-4) with explicit instructions to critically revise AI outputs, while the control group relied on dictionaries and personal knowledge.

Table 2
Qualitative Statistics’ Analytical Dimensions

Analytical Dimension	Interview Questions
General Experience Language	1. How would you describe your overall experience using ChatGPT in your translation lessons?
	2. Do you think that ChatGPT could help you improve your vocabulary and grammatical accuracy in translation?
	3. Did you notice any errors or unnatural phrasing generated by ChatGPT? How did you deal with them?
Intercultural	4. How well do you think ChatGPT understands Vietnamese cultural nuances or idiomatic expressions when translating?

Thematic	5. How does ChatGPT help you understand or express complex ideas from the source text?
	6. Do you feel your translations using ChatGPT better preserve the message and logical flow of the original text compared to manual translation? Why or why not?
Strategic	7. When ChatGPT gives you a translation, what steps do you usually take afterward?
	8. How do you decide whether ChatGPT's output is acceptable or requires major revision?
Technological	9. How did you learn to use ChatGPT effectively for translation?
	10. Do you feel that ChatGPT could replace some parts of the translation process, or should it remain a support tool? Why?
Reflective Wrap-Up	11. Looking back, how has ChatGPT affected your learning of translation skills? If your teachers integrate ChatGPT-based activities into translation courses, what would you like those activities to focus on?

Results

Quantitative Statistical Findings

The results of the quantitative statistical analysis reveal a coherent pattern that illuminates the differential impact of ChatGPT-assisted translation training across four EMT dimensions. As can be seen, MAN group outperformed GPT group in terms of posttest scores ($M=3.93$ as opposed to $M=3.60$ respectively). Based on the Std. Deviation (SD), the former showed a greater variability (0.5) while the latter reported a smaller figure of 0.25, approximately half as much. Accordingly, the more significant the variation was, the more marked difference students' ability showed. In contrast, the smaller gap indicated the more consistent performance.

Table 3

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	MAN-Pretest	3.00	15	.000	.000
	GPT-Pretest	3.07	15	.258	.067
Pair 2	MAN Posttest	3.60	15	.507	.131
	GPT Posttest	3.93	15	.258	.067

Table 4 shows that the pre-test correlation between MAN group and GPT group could not be computed because the MAN's pre-test scores had no variation ($SD = 0$), making it impossible to assess the relationship between the two variables. In contrast, the post-test showed a slightly different pattern ($r = 0.327$; $p = 0.234$). This means the post-test scores of both groups of students do not depend on each other; in other words, the correlation between them was low and not statistically significant.

Table 4

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	MAN-Pretest & GPT-Pretest	15	.	.
Pair 2	MAN Posttest & GPT Posttest	15	.327	.234

As can be seen from Table 5, there was no statistically significant difference between the surveyed groups at the pre-test stage, as indicated by a minor mean difference of -0.067 , $t(14) = -1.000$, $p = .334$, with the 95% confidence interval $[-0.210, 0.076]$.

The result confirms that the two groups were equivalent before the intervention. In contrast, a significant difference emerged at the post-test, where the mean difference increased to -0.333 , $t(14)$ is -2.646 , p equals $.019$, and the confidence interval $[-0.604, -0.063]$ did not include zero.

These findings indicate that the GPTgroup scored significantly higher than the MAN group after the intervention, with GPTs consistently outperforming MANs at the post-test stage.

Table 5
Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	SD	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	MAN-Pretest - GPT-Pretest	-.067	.258	.067	-.210	.076	-1.000	14	.334
Pair 2	MAN Posttest - GPT Posttest	-.333	.488	.126	-.604	-.063	-2.646	14	.019

To examine how the two groups performed in each dimension in general and in 15 specific sub-competences in particular, have a quick look at Table 6.

Regarding the pre-test results, the mean scores of all 15 sub-competences of both groups were not much different. The biggest disparity can be seen in Sub-2 (lexical equivalents) and Sub-5 (textual equivalents) with 0.2 equally. In these two aspects, the experimental group performed better. The smallest difference was in Sub-3, stylistic & semantic appropriateness, and Sub-8, pragmatic equivalence, with the figure of 0.01 each. There are three sub-competences in which the control group did better than the experimental group. They are Sub-1, Sub-3, and Sub-7.

Turning to the posttests, we can realize significant differences in the performance of the two surveyed groups. Obviously, all of the students made progress after this 15-week study; however, those who were allowed to use ChatGPT in their translation lessons and tests made greater strides than those who did the translation manually.

The most noticeable dissimilarity is seen in the translation of stylistic and semantic refinement, with the GPTgroup having a mean score of 3.98, but the MAN group achieved 3.42, even though this was not the aspect that had the most noticeable improvement. This is followed by sub-15, text typology, or the ability to classify texts of different types for translation facilitation, with a difference of 0.52, in which the

score of the experimental group was 3.97, and that of the control group was 3.45. Grammar accuracy and thematic/domain knowledge are also among the sub-competences with the most significant difference, with 0.36 and 0.24, respectively.

Based on the statistical findings, the sub-competences that witnessed the smallest differences are Sub-7, the ability to render connotational meanings, and Sub-11, subject-matter understanding. Both of them represented a gap of 0.16. Sub-8, the translation of pragmatic equivalence from source texts to target texts, also exhibited a minimal gap of 0.18.

Table 6
Pretest and Posttest Mean Score Comparison

Analytical Dimension	Sub-competence	MAN-Pretest	GPT-Pretest	MAN-Posttest	GPT-Posttest
Language Competence	Sub1	3.25	3.23	3.92	4.28
	Sub2	3.28	3.48	3.77	4.12
	Sub3	2.98	2.97	3.42	3.98
	Sub4	3.4	3.57	3.68	4.07
	Sub5	3.17	3.37	3.55	4.07
Intercultural Competence	Sub6	3.18	3.23	4	3.61
	Sub7	3.23	3.18	3.72	3.88
	Sub8	3.02	3.03	3.35	3.07
	Sub9	3.17	3.25	3.52	3.77
	Sub10	3.17	3.3	3.47	3.68
Thematic Competence	Sub11	3.18	3.22	4.03	4.19
	Sub12	3.13	3.1	3.73	3.97
	Sub13	3.07	3.05	3.47	3.75
	Sub14	3.05	3.12	3.37	3.87
	Sub15	3.15	3.27	3.45	3.97

As for the improvements of the two groups, the supported students showed substantial increases in all 15 sub-competences. Their grammar accuracy received the most marked increase of 1.05. Second came the stylistical and semantic appropriateness with an increase of 1.01. Sub-8, pragmatic equivalence, showed almost no change, remaining at around 3.0.

Positive changes could also be observed in the control group with all 15 sub-competences; however, this group's improvements were not comparable to the experimental group. What stands out from the table is that score gains in the former group were not parallel to the gains of those with ChatGPT support. Sub-11, students' showing subject-matter understanding, and Sub-6, students' rendering cultural nuances in translation, were the two best-performed skills with 0.85 and 0.82, respectively. Same as the GPTgroup, which included sub1, grammatical accuracy, in the three best sub-competences, the MAN group indicated considerable grammar enhancement, but its data was only half that of the GPT assisted participants. The

three least improved sub-competences of the MAN group are Sub-14 (+0.32), Sub-8 (+0.33), and Sub-15 (+0.30).

Discussion

ChatGPT and Translator's Language Competence

Comparing the performance of both groups' language competence, although the two groups, after the intervention, showed the strongest achievement in grammar accuracy, the experimental students had a greater improvement than those doing translation tasks manually in this aspect (1.05 as opposed to 0.67). A similar pattern can be found in the production of stylistic and semantic appropriateness, but in this regard, the GPTgroup even gained bigger achievements than the MAN group (1.01 in comparison with 0.44). There were positive changes in the use of lexical equivalents and textual equivalents, but the differences between the two categories of respondents were not as significant as the first two mentioned grammatical features. Both groups showed the most model changes in cohesive devices' usage, and the discrepancy by group was also the least significant. From the statistical findings, the GPTgroup could perform better in all subcompetences, but what is most remarkable is that, contrary to the general belief, it could surprisingly outperform in the respects that require more sensibility, such as grammatical range and stylistic scale.

Responses from the five interviewees to some extent echo the quantitative findings. IN1 agreed that this tool supported them to better their vocabulary and grammar in use because she felt that when using ChatGPT, her translation is "*free of grammatical errors*", adding that "*some words that I don't know, ChatGPT can help.*" This opinion was in line with IN2's explanation, which confirmed that she "*doesn't have to use a dictionary because ChatGPT gives the words very quickly*".

The results of this study confirm that ChatGPT exerts a measurable and differentiated influence on students' translation competence when assessed through the EMT framework. Quantitative analysis revealed statistically significant gains in language competence. Such achievements were in alignment with the findings from Deng (2024) who stated ChatGPT translators outperformed human translators in language use, Dinh (2025) who confirmed that with this app, students achieved higher surface accuracy and stylistic appropriateness, Zhao (2025) who believed in the outstanding ability of ChatGPT to work with simple sentences, or Haryanti (2024) who reaffirmed the accuracy of translation conducted by this tool.

However, there existed a paradox as both IN3 and IN4 were not quite sure about the positive impact of ChatGPT on their grammar competence. While the former said "*I don't know. Because my English is still not good. Even after I use ChatGPT, I just use it so that I can do the translation fast. And because my teacher asked us, and I need my translation good*", the latter admitted that "*Um, actually, it can help, but, um, I – I'm not quite sure whether it helps to improve or it helps me to do the translation (laughs) practices. Um, do you understand me? I mean that I can do the translation*

very quickly, but, um, after that, I do not remember much. So I'm ... I don't know whether it helps me to learn or it helps me to do the translation test." IN5 even rejected that his grammatical accuracy can be really improved, though he said that his vocabulary benefited from the use of ChatGPT. He added that his grammar was not good and he did not *"read it again, and ...do not try to learn and analyze the sentence."* Such students' confession makes us reconsider the actual impacts of ChatGPT incorporation or at least try to work out the best possible methods to make use of ChatGPT's intervention in EFL classes or translation training courses.

ChatGPT and Translator's Intercultural Competence

What should be first noticed is that there was also no difference between the two groups, both in pretest scores and posttest results, when the intercultural competence was evaluated. As predicted, the MAN group scored higher in rendering cultural nuances and pragmatic equivalence. In this dimension, the MAN group elucidated the biggest improvement in sub 6, which showed their ability to handle subtle cultural nuances in translation (an increase of 0.82 mean score). However, surprisingly, the supported group still got better marks in other cultural aspects, including working with the connotational meanings, realizing contextual equivalence, and exhibiting the adaptation of tone, register, and communicative intent. This indicates that ChatGPT has been able to recognize basic cultural concepts or meanings; nonetheless, it could not yet feature the in-depth or subtle nuances due to the fact that there was almost no change in the score gained by GPTgroup in capturing pragmatic equivalence in translation, though for this domain, the MAN group gained a considerable score of more than 0.3.

These findings did not align with Nguyen (2024), who did not agree that ChatGPT was a tool to translate cultural concepts or the metaphorical meanings of words, Chen and Lin (2025), recognizing ChatGPT's failure to render culturally nuanced content, or Khoshafah (2021) who admitted this AI instrument could not help with "jargon and terms". Increases in the mean score of almost all cultural sub-competences of the GPTgroup also went against what Yusuf and Sivanadhan (2025) concluded "challenges related to idiomatic expressions [...] remain" (p.251) despite substantial improvements in the use of AI pre-trained language models.

The quantitative findings also were not in line with the qualitative results because when questioned about the translation of culture-bound or idiomatic language using ChatGPT, interviewees voiced skepticism. Their responses were filled with key phrases such as *"cannot understand culture"* (IN1), *"Western style"* (IN2), and *"funny translation"* (IN3). IN1 laughed and admitted that ChatGPT found it hard to work with *"underlying meanings sometimes"* and that ChatGPT translated Vietnamese proper names *"in a very funny way"*. IN4 strengthened such a point of view when adding that *"I usually have to double-check, especially when it comes to proper names or cultural concepts."* IN5 elaborated further: *"ChatGPT just uses the Western ways, not the*

Vietnamese way. Even the way Vietnamese people express the sentence, it translates differently.”

That ChatGPT can work effectively, to some extent, in a culturally rich context is a promising prospect and this can possibly be attributed to the better database of this AI tool. In other words, ChatGPT or other AI tools have been constantly updated, making it increasingly human-like in EFL teaching and learning.

ChatGPT and Translator's Thematic Competence

The thematic competence dimension reveals another aspect of the intervention. Although both groups progressed and the ChatGPT users could perform better than the manual translators, changes in all five sub-competences of the GPTgroup were more consistent, ranging from 0.7 to 0.9. However, the gain-in scores of the MAN group varied considerably from 0.3 (Sub15) to 0.85 (Sub11). This indicates the variations in human perception and ability, but the consistency in AI-driven translation.

Results from the tests were further consolidated by the qualitative data taken from the interviewees. IN1, IN2, and IN3 agreed that they could count on ChatGPT to build domain-specific knowledge. IN1 said she could find what she needed from ChatGPT, while IN2 was happy to “*find accurate terms*” and considered this chatbot an important source of knowledge.

Both the quantitative and qualitative findings agreed with Zhao (2025) and Zhou (2025), who noted that generative models prioritize fluency over factual accuracy and require human oversight to preserve conceptual integrity. To justify the better performance of the students in this domain, ChatGPT's access to extensive lexical databases should be the main factor.

However, the research findings regarding this dimension did not receive the total agreement from the respondents as IN4 and IN5 stated that they had a habit of double-checking their translation when they used ChatGPT for some technical documents as they “*doubt the accuracy of terminological translation*” as IN4 stated or the expressions sometimes “*not quite appropriate*” (IN5). The students' responses were not in line with the substantial increases in the score of sub-competences in this respect. This fact again echoes our concern about the real efficiency of students using ChatGPT.

ChatGPT and Translator's Strategic and Technological Competence

Unlike intercultural and thematic dimensions, strategic competence achieved a general agreement among the interviewees. Findings displayed both statistical and experiential growth in this competence. Respondents described rereading and adjusting the AI's output, using verbs such as “check,” “compare,” and “correct.” IN1 reported, “*Normally, I check it, read it again, mainly for vocabulary or idiomatic*

meaning,” while IN4 explained, “*I read it again, check, and correct the problems. Normally, I accept it one hundred percent, only minor revision.*” The findings are consistent with Jiang et al. (2024) and Nguyen (2024), indicating ChatGPT as a diagnostic tool that encouraged reflection through contrastive learning, which can be considered as an important aspect of learner autonomy.

The qualitative findings themselves expose a paradox. While students agreed that they developed the habit of translation revision, a sign of autonomous learning, or they were “*no longer worried about translation tasks*” (IN2), some confessed to “*feel lazy*” (IN1), “*rely too much*” (IN4) on the AI, or “*my brain doesn't have to work*” (IN3). The results resonate Fan et al. (2023), who warn that uncritical dependence diminishes intellectual effort and ethical responsibility. This ambivalence reveals that ChatGPT can both enhance and undermine autonomy depending on instructional framing. Therefore, the assertion that learn autonomy had been fostered and boosted thanks to ChatGPT, as what Nguyen (2024) and Chen and Liu (2025) confirmed, requires further empirical validation.

The final dimension, the technological, was responded to with not much difference. All of the interviewed students agreed that they had little to do with the technological requirements of using ChatGPT. “*Just install it and use*” (IN1), “*it takes a single minute to have this app*” (IN3), “*easy and simple*” (IN4), or “*phone, just need a phone and I can use it right in class*” (IN5). The answers mean that students do not need to spend much time working on this technological application, and they appear not to require instructions or guidance.

Conclusion


This study examined the effects of integrating ChatGPT into EFL translation training. It compared the performance and perceptions of students who were required to complete their translation tasks with and without ChatGPT assistance. Students' translation assessment was based on four EMT dimensions, which are linguistic, intercultural, thematic, and strategic competences.

The findings indicate that ChatGPT is an effective facilitator in a translator course. Quantitatively, the experimental group reported improvements in their language and strategic competence. These results can be attributed to ChatGPT's capacity to enhance students' grammatical accuracy and language use. In addition, through iterative comparison and feedback, their strategic competence was also developed. That no significant achievement in intercultural competence was reported is understandable because culture-bound texts are known to lie still beyond machine translation's reach. The study also found that students' cultural competence got better, as shown via considerable mean score increases. This finding reflects the potential of ChatGPT to evolve and, in the future, to possibly surpass certain human capacities. However, there still remains disagreement as to whether this highly evolved translation bot can foster learner autonomy and improve translation competence is still

an open question. The paradoxical findings in students' thematic competence, together with the quantitative gains but qualitative inconsistencies in their strategic competence, are something worth considering.

Echoing the research findings, I believe that the incorporation of ChatGPT into translation training programs should be approached with prudence. The first reason is the increases in test scores may not reflect the real improvements in learners' language and translation efficiency. Second, the possibility that students cannot keep their learning momentum needs verification. As such, ChatGPT should only be positioned as a supportive rather than a substitutive tool in curriculum design and development.

ORCID

 <https://orcid.org/0009-0007-1342-4231>

Publisher's Note

The claims, arguments, and counter-arguments made in this article are exclusively those of the contributing authors. Hence, they do not necessarily represent the viewpoints of the authors' affiliated institutions, or EUROKD as the publisher, the editors and the reviewers of the article.

Acknowledgements

Not applicable.

Funding

This research received no grant from any funding agency.

CRedit Authorship Contribution Statement

Tran Thi Ngoc Lien: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project Administration

Generative AI Use Disclosure Statement

The author used ChatGPT to assist with language editing and phrasing. All ideas, analyses, and interpretations were developed independently by the author.

Ethics Declarations

World Medical Association (WMA) Declaration of Helsinki–Ethical Principles for Medical Research Involving Human Participants

This study was conducted in accordance with the ethical standards of the Declaration of Helsinki. Participation was entirely voluntary and consent forms were obtained from all participants. All data were collected anonymously and used solely for research purposes.

Competing Interests

The author declares no competing interests

Data Availability

The data presented in this study are available from the corresponding author upon reasonable request.

References

- Alosaimi, B. A. A., & Alawad, N. A. (2024). Evaluation of the translation of separable phrasal verbs generated by ChatGPT. *Arab World English Journal (Special Issue on ChatGPT)*, 282–291. <https://doi.org/10.24093/awej/ChatGPT.19>
- Angelone, E. (2010). Uncertainty, uncertainty management and metacognitive problem solving in the translation task. *Across Languages and Cultures*, 11(1), 17–40. <https://doi.org/10.1075/ata.xv.03ang>
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford University Press
- Banat, M., & Adla, A. Y. (2023). Exploring the effectiveness of GPT-3 in translating specialized religious text from Arabic to English: A comparative study with humans. *Journal of Translation and Language Studies*, 4(2), 1–23. <https://doi.org/10.48185/jtls.v4i2.762>
- Bell, T.R. (1991). *Translation and Translating: Theory and Practice*. Longman
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589–597. <https://doi.org/10.1080/2159676X.2019.1628806>
- Bui, T. K., & Tran, A. V. (2024). AI-assisted learning applications in English courses in Vietnam. *Journal of Applied Learning & Teaching*, 7(2), 99–116. <https://doi.org/10.37074/jalt.2024.7.2.34>
- Chen, P. (2024). The impact of generative AI on the role of translators and its implications for translation education. *Education Insights*, 1(2), 1–10. <https://soapubs.com/index.php/EI>
- Chen, S., & Lin, Y. (2025). A multidimensional comparison of ChatGPT, Google Translate, and DeepL in Chinese tourism text translation: Fidelity, fluency, cultural sensitivity, and persuasiveness. *Frontiers in Artificial Intelligence*, 8, 1–15. <https://doi.org/10.3389/frai.2025.1619489>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
- Dinh, C. T., et al. (2025). EFL students' perspectives on ChatGPT in translation: Exploring AI assistance, motivation, and learning outcomes. *Electronic Journal of e-Learning*, 23(2), 99–116. <https://doi.org/10.34190/ejel.23.2.4006>
- Esfandiari, R. M., Shokrpour, N., & Rahimi, F. (2019). An evaluation of the EMT: Compatibility with professional translators' needs. *Cogent Arts & Humanities*, 6(1), 1601055. <https://doi.org/10.1080/23311983.2019.1601055>
- Fan, P., Gong, H., & Gong, X. (2023). The application of ChatGPT in translation teaching: Changes, challenges, and responses. *International Journal of Education and Humanities*, 11(2), 49–52. <https://doi.org/10.54097/ijeh.v11i2.13530>
- Ghassemiazghandi, M. (2024). An evaluation of ChatGPT's translation accuracy using BLEU score. *Theory and Practice in Language Studies*, 14(4), 985–994. <https://doi.org/10.17507/tpls.1404.07>
- Göpferich, S. (2009). Towards a model of translation competence and its acquisition: The longitudinal study TransComp. University of Graz. <https://gams.uni-graz.at/o:tc-095-187>
- Hatim, B. & Mason, I. (1997). *The translator as communicator*. Routledge.
- Huang, L., Li, M., & Lee, C. (2020). Neural machine translation for low-resource languages: A review. *Journal of Artificial Intelligence Research*, 69, 593–617. <https://doi.org/10.1613/jair.1.11873>
- Jalambo, M., Aladini, A., & Amer, M. M. (2023). Evaluating the MA translation programme at IUG in light of the EMT competences. *Multidisciplinary Science Journal*, 5, e2023056. <https://doi.org/10.31893/multiscience.2023056>

- Jiang, L., Jiang, Y., & Han, L. (2024). The potential of ChatGPT in translation evaluation: A case study of the Chinese-Portuguese machine translation. *Cadernos De Tradução, 44*(1), 1–22. <https://doi.org/10.5007/2175-7968.2024.e98613>
- Karataş, F., Yaşar, F., & Gınyel, F. O. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies, 29*(15), 19343–19366. <https://doi.org/10.1007/s10639-024-12574-6>
- Kelly, D. (2005). *A handbook for translator trainers: A guide to reflective practice*. St. Jerome Publishing. <https://www.researchgate.net/publication/392809864>
- Khoshafah, F. (2023). ChatGPT for Arabic–English translation: Evaluating accuracy. *Research Square, 1–18*. <https://doi.org/10.21203/rs.3.rs-2814154/v2>
- Kim, N. (2018). Chatbots and Korean EFL students' English vocabulary learning. *Journal of Digital Convergence, 16*(2), 1–7. <https://doi.org/10.14400/JDC.2018.16.2.001>
- Kiraly, D. (2014). A social constructivist approach to translator education: Empowerment from theory to practice. *Meta: Journal des Traducteurs, 59*(1), 76–93. <https://doi.org/10.5007/5899>
- Leighton, J. P., & Gierl, M. J. (Eds.). (2007). *Cognitive diagnostic assessment for education: Theory and applications*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511611186>
- Liu, C., Wei, Z., & Wang, Y. (2021). Advancements in neural machine translation: Techniques and challenges. *IEEE Transactions on Neural Networks and Learning Systems, 34*(4), 567–582. <https://doi.org/10.1109/TNNLS.2021.3060678>
- Nida, E. (1976). A framework for the analysis and evaluation of theories of translation. In R. W. Brislin, *Translation: Applications and Research* (pp. 47-91). Garner Press.
- Nguyen, T. T. H. (2024). Integrating ChatGPT in translation learning. In *Annual SEAAIR Conference Proceedings: SEAAIR 2024 - The 24th Annual Conference (Vol. 4, pp. 603–610)*. SEAAIR. <http://www.seairweb.info/conference/2024Proceedings.pdf>
- Nguyen, T. N. N. (2025). Enhancing competencies through AI technology and translation integration in the bilingual wall newspaper project. *Computer-Assisted Language Learning Electronic Journal (CALL-EJ), 26*(3), 1–23. <https://doi.org/10.54855/callegj.252631>
- Plaza-Lara, C. (2022). Competences of translation project managers from the academic perspective: Analysis of EMT programmes. *The Interpreter and Translator Trainer, 16*(2), 203–223. <https://doi.org/10.1080/1750399X.2021.1987085>
- Pym, A. (2003). Redefining translation competence in an electronic age. *Meta, 48*(4), 481–497. <https://doi.org/10.7202/008533ar>
- Rothwell, A., & Svoboda, T. (2019). Tracking translator training in tools and technologies: Findings of the EMT survey 2017. *The Journal of Specialised Translation, 32*(1), 26–60. https://commission.europa.eu/document/download/b482a2c0-42df-4291-8bf8-923922ddc6e1_en?filename=emt_competence_fw_k_2022_en.pdf
- Sahari, Y., Al-Kadi, A. M. T., & Ali, J. K. M. (2023). A cross-sectional study of ChatGPT in translation: Magnitude of use, attitudes, and uncertainties. *Journal of Psycholinguistic Research, 52*(6), 2937–2954. <https://doi.org/10.1007/s10936-023-10031-y>
- Schäffner, C., & Adab, B. (2000). *Developing translation competence*. John Benjamins.
- Schäffner, C. (2012). Translation competence: Training for the real world. In S. Hubscher-Davidson & M. Borodo (Eds.), *Global trends in translator and interpreter training: Mediation and culture* (pp. 30–44). Continuum International Publishing Group.
- Sun, R. (2024). Evaluating the translation accuracy of ChatGPT and DeepL through the lens of implied subjects. *AWEJ for Translation & Literary Studies, 8*(4), 41–53. <https://doi.org/10.24093/awejtls/vol8no4.5>
- Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages, 8*(3), 1–12. <https://doi.org/10.3390/languages8030212>
- Yusuf, R., & Sivanadhan, I. (2025). The role of artificial intelligence in enhancing language translation and cross-cultural communication. *Malaysian Journal of Social Sciences and Humanities (MJSSH), 9*(4), 245–256. <https://doi.org/10.33306/mjssh/379>

Tran Thi Ngoc Lien

- Zhao, S. (2025). Applications and challenges of ChatGPT in translation: A multidimensional exploration of theory and practice. *Sino-US English Teaching*, 22(2), 35–41. <https://doi.org/10.17265/1539-8072/2025.02.01>
- Zhou, L. (2025). Evaluating ChatGPT's Chinese–English translation quality of tender documents: A research of MTPE with MQM scoring models. In *Proceedings of the 4th International Conference on Literature, Language, and Culture Development* (pp. 53–62). <https://doi.org/10.54254/2753-7064/61/2025.20639>
- Will, W. (1976). A framework for the analysis and evaluation of theories of translation. In R. W. Brisilin, *Translation: Applications and Research* (pp.117-137). Garner Press.