Translation Studies, Winter 2026, Vol. 23, No. 92, pp. 21–36: **Scientific Research Paper**
DOR: https://dor.isc.ac/dor/20.1001.1.17350212.1404.23.4.3.9

Impact of Calibrated AI-Prompts on Translation Quality: Insights from Cognitive Modifiability Perspectives[8]

Mahboobeh Khosravani[9]

Abstract

This study investigated whether AI prompts calibrated to learners' zones of proximal development (ZPD) improve translation quality more effectively than unguided AI assistance or conventional teacher feedback. Using a sequential mixed-methods, quasi-experimental design with 32 Iranian undergraduate translation students, a qualitative diagnostic phase identified four error types (lexico-semantic, syntactic-grammatical, pragmatic, meta-functional) and produced a ZPD-calibrated mediational inventory of graduated implicit-to-explicit prompts. In the quantitative phase, participants were assigned to three conditions: ZPD-tuned AI mediation (ChatGPT with calibrated prompts), normative AI mediation (standard ChatGPT use), or teacher feedback. One-way ANOVA showed a significant effect of mediation type, $F(2, 29) = 33.574$, $p<.001$. The ZPD-tuned AI group significantly outperformed both the normative AI group ($MD = 5.15$, $p<.001$) and the teacher feedback group ($MD = 4.79$, $p<.001$), with no significant difference between the latter two. Findings indicate that developmentally calibrated, scaffolded AI prompts produce superior translation outcomes compared to unguided AI access or conventional instruction, highlighting the value of integrating ZPD-aware prompt engineering into translator education.

Keywords: AI-mediated translation, Cognitive modifiability, Mediation, Prompt engineering, Translation quality

8. This paper was received on 26.12.2025 and approved on 17.01.2026.

9. Lecturer, Department of English Language, Faculty of Literature and Humanities, University of Birjand, Birjand, Iran; email: m.khosravani@birjand.ac.ir

1. Introduction

Machine translation, based on large language models, has become one of the most well-established areas of research in linguistic analysis in current decade and thus, due to its theoretical and practical expansion, has attracted many researchers in academia and industry (Jiao et al., 2024). Many studies (e.g., Wang, Zhang, 2025; Wu & Ding, 2025; Gao, Wang, & Hou, 2023), consistently demonstrate that linguistically and culturally informed prompt engineering enhances AI-generated translation quality in terms of accuracy, fluency, and contextual appropriateness. However, this line of research has predominantly adopted a model-centered orientation, focusing on optimizing AI outputs while largely overlooking the cognitive characteristics and developmental needs of human translators who interact with these systems. In contrast, research grounded in sociocultural theory (Vygotsky, 1978), dynamic assessment (Poehner, 2008), and structural cognitive modifiability (Feuerstein et al., 2002) has shown that linguistic performance improves when instructional prompts are responsively calibrated to learners' errors and their ZPD through graduated mediation. Building on these insights, the present study re-conceptualizes AI not merely as a translation engine but as a mediational artifact capable of delivering graduated, learner-responsive prompts. Specifically, it investigates whether AI-generated prompts, systematically tuned to individual learners' ZPD as identified through error-based diagnostic mediation, can significantly enhance the quality of human translation output in AI-mediated translation tasks. Accordingly, this study is guided by the following research questions: a) what are the dominant categories and distributions of translation errors identified during the diagnostic phase under human-mediated conditions? And b) to what extent does the type of mediation (ZPD-calibrated AI, normative AI, or teacher-mediated feedback) differentially affect learners' translation quality?

2. Literature Review

The reviewed literature highlights consistent gains from AI use in translation while simultaneously revealing a critical gap regarding the role of learner-responsive, developmentally informed prompting. Khosravani et al. (2025) examined the comparative effectiveness of AI-assisted, AI-human revised, and human-scaffolded translation in English for Specific Purpose (ESP) instruction. Using a mixed-methods, posttest-only quasi-experimental design, and the study involved 46 Iranian postgraduate ESP students divided into three instructional conditions. Translation quality was analyzed through rubric-based quantitative measures and qualitative

error analysis. The findings showed that the AI-only condition exhibited substantial lexico-semantic, syntactic, and discourse-level deficiencies, whereas conditions involving human mediation significantly outperformed the AI-only group. These results underscore that AI feedback in isolation is insufficient for high-quality translation and point to the central role of mediation. However, the study does not explore whether AI itself can be designed to deliver developmentally responsive mediation rather than static feedback.

Focusing on translator-AI interaction from a learner perspective, Zhang and Doherty (2025) investigated postgraduate translation students' use of generative AI tools, with particular emphasis on prompting behaviors. Analyzing 983 prompts collected over eight months through thematic and discourse analysis, the study found that students primarily relied on GenAI for transfer and revision, while rarely using it for analytical or documentary purposes. Prompts were largely intuitive, underspecified, and weakly contextualized, indicating limited awareness of effective prompt engineering. While this study highlights a pedagogical gap in students' prompting competence, it does not examine how prompts could be systematically scaffolded or calibrated to learners' cognitive needs. From a model-centered perspective, Mondshine et al., (2025) explored how selective pre-translation of prompt components affects multilingual prompting performance in large language models. Through a large-scale experimental design covering 35 languages, the study demonstrated that selectively translating prompt elements (e.g., instructions or examples) consistently outperforms full pre-translation and direct inference. Although the findings offer a robust strategy for improving model performance, the study prioritizes optimization of AI output rather than the cognitive engagement or developmental responsiveness of human users. Similarly, Zhang and Fan (2025) investigated the effects of prompt strategies on ChatGPT translation quality for specialized texts across news, legal, and medical domains. Using a controlled experimental design, the study showed that genre-based prompts yield the most consistent improvements, while register and audience prompts are particularly important for technical domains. Despite providing valuable insights into domain-sensitive prompt design, the study remains focused on text-level and genre-level optimization, without considering individual learner differences or adaptive mediation.

Complementing these task-level studies, broader trends in translation research indicate a growing emphasis on AI-mediated translation and human–machine collaboration. Bibliometric analyses reveal that neural machine translation, human

post-editing, and AI-assisted tools dominate contemporary research agendas (Mohsen, Althebi, & Albahooth, 2023). Empirical evidence from live-streaming contexts further suggests that machine translation combined with human revision generally outperforms computer-assisted simultaneous interpreting in accuracy and efficiency, though outcomes vary by text type and evaluative framework (Liang & Lu, 2025). From an educational standpoint, research shows that while translation students frequently use AI tools and express confidence in post-editing, they often lack ethical awareness and strategic competence, underscoring the need for principled integration of AI literacy in translator education (Zhang & Doherty, 2025). Taken together, these studies demonstrate that AI-assisted translation benefits substantially from human involvement and informed prompt design. However, existing research has largely approached prompting from either a model-centered optimization perspective or a pedagogical skills perspective, leaving unexplored the potential of AI prompts that are systematically tuned to learners' zone of proximal development. Notably absent is empirical research examining whether AI-generated prompts, designed as graduated and error-responsive mediation, can enhance human translation performance. Addressing this gap, the present study investigates the impact of ZPD-tuned AI prompts on translation quality within AI-mediated translation tasks.

3. Methodology

3.1. Research Design

The present study adopted a mixed-methods, quasi-experimental design consisting of an initial diagnostic qualitative phase followed by a quantitative experimental phase. This design was informed by principles of socio-cultural theory and structural cognitive modifiability, allowing learner responsiveness to mediation to guide the subsequent design of AI-based prompts. In the first phase, a diagnostic dynamic assessment procedure was implemented to identify learners' translation difficulties and responsiveness to mediation in the absence of AI. Based on the mediation patterns and error profiles observed during this phase, a typology of mediation types and responsiveness moves was extracted. This typology then served as the basis for designing a graduated AI-prompt inventory implemented through ChatGPT. In the second phase, the effectiveness of ZPD-tuned AI mediation was empirically tested through comparison with two alternative instructional conditions, i.e., AI mediation using regular, non-graduated ChatGPT prompts, and traditional teacher-mediated feedback provided in a normative classroom setting. Translation

performance across the three conditions was compared quantitatively to examine the impact of developmentally tuned AI prompts on translation quality.

3.2. Participants

The participants were 32 intermediate-level undergraduate translation students enrolled in a translation course at university of Birjand in Iran. They constituted an intact group, selected through convenience sampling due to institutional constraints. Participants had previously completed foundational courses in translation theory and practice and shared a relatively homogeneous proficiency level as determined by departmental placement criteria. Given the intact-group design, participants were randomly assigned to three experimental conditions to minimize selection bias.

3.3. Translation Rating Rubric

All translations were evaluated using an analytic scoring rubric developed in alignment with the error domains identified in the diagnostic phase. The rubric comprised four weighted dimensions: (a) lexico-semantic accuracy, (b) syntactic-grammatical configuration, (c) pragmatic/register appropriateness in political discourse, and (d) meta-functional alignment (including preservation of process type, participant roles, modality, and agency). Each dimension was rated on a 5-point scale ranging from 1 (serious distortion of meaning/function) to 5 (accurate and functionally equivalent rendering). Two trained raters with graduate-level expertise in translation studies independently scored all scripts. Prior to scoring, raters underwent calibration sessions using benchmark translations to ensure shared interpretation of rubric descriptors. Inter-rater reliability was calculated using intra-class correlation coefficients, and discrepancies exceeding one scale point were resolved through discussion to reach consensus. The composite score for each participant was computed by summing the four dimensions, yielding an overall translation quality index (ranging 4–20) for subsequent statistical analysis.

4. Results & Discussion

This section presents the findings of the study in two phases. First, results from the qualitative diagnostic phase are reported, focusing on the identification of translation error patterns and learners' responsiveness to mediation. Second, quantitative results are presented, comparing translation quality across the ZPD-tuned AI mediation, normative AI mediation, and teacher-mediated feedback conditions.

4.1. Qualitative Phase

4.1.1. Diagnostic Stage

The qualitative phase aimed to identify recurrent translation error patterns that emerged during the human-mediated diagnostic assessment. Learners' translations of a political text adopted from mainstream technical university-level course books in the country, were analyzed through iterative coding, combining semantic, syntactic, pragmatic, and systemic functional criteria. Error identification was guided by two principles: (a) deviation from target-language norms in political discourse, and (b) distortion of experiential meaning, participant configuration, or interpersonal stance. The analysis yielded four major categories of errors: Lexico-Semantic Errors, Syntactic-Grammatical Errors, Pragmatic Errors, and Meta-Functional Errors. Within each category, several sub-categories were identified based on recurring patterns. Table 4.1 presents the typology of errors along with representative examples extracted from learners' translations.

Table 4.1. *Typology of Translation Errors in Diagnostic Phase*

Error Category	Error Sub-Category	Examples
Lexico-Semantic Errors	Non-Equivalent Lexical Item	• The law was ***passed*** • قانون رد شد (صحیح: تصویب شد)
	Meaning-Overgeneralization	• Demanded ***accountability*** • خواستار پذیرش مسئولیت شد (صحیح: پاسخگویی)
	Semantic Narrowing	• Nuclear ***Proliferation*** • گسترش هسته‌ای (ترجمهٔ صحیح: اشاعهٔ هسته‌ای)
	Metaphorical Misinterpretation	• It was a ***bipartisan*** measure • این اقدام ناشی از دودستگی بود (صحیح:این اقدام با توافق هر دو گروه بود)
	Inaccurate Terminology	• The government ***ratified*** the treaty • دولت معاهده را تصویب کرد (صحیح: تصویب نهایی)
	Terminology Confusion	• Electoral college

		• دانشگاه انتخاباتی (صحیح: مجمع گزینندگان)
	Polysemic Slip	• The party ***holds*** the majority • حزب ، اکثریت را نگه می‌دارد (صحیح: در اختیار دارد)
	Literal Rendering	• ***Power-sharing*** agreement • توافق تقسیم قدرت (صحیح: مشارکت در قدرت)
	Colloquial mismatch	• The tension ***escalated*** • تنش‌ها بیشتر شد (صحیح: تشدید شد)
Syntactic-Grammatical Errors	Role Reversal	• The resolution, ***vetoed by*** the president, sparked public outrage • قطع‌نامه‌ای که *رئیس جمهور را* سلب اختیار کرد باعث شورش شد (صحیح: رئیس جمهور نقش فاعل جملهٔ مجهول)
	Wrong Phrase Ordering	• The ***newly proposed economic sanction*** • جدیدا تحریم‌های اقتصادی پیشنهاد شده (صحیح: تحریم‌های اقتصادی تازه پیشنهاد شده)
	Future-in-past	• It was expected that negotiation ***would resume*** • انتظار می‌رفت که مذاکرات از سر گرفته شود (صحیح: از سر گرفته خواهد شد)
	Agreement error	• The delegation of diplomats ***was delayed*** • هیات دیپلمات‌ها تاخیر داشتند (صحیح: تاخیر داشت)
Pragmatic Errors	Register Misalignment	• We express our deep concern • ما عمیقا نگران هستیم (صحیح: نگرانی خود را ابراز می‌کنیم)
	Register Downgrading	• Legislative Framework • قوانین (صحیح: چارچوب قانونی)

	Agent (Participant) Deletion	• ***The country*** imposed new sanctions on the oil sector • تحریم‌های جدیدی بر بخش نفت اعمال شد (صحیح: اضافه کردن عامل دستوری)
Meta-Functional Errors	Verb Group shift (mental-material)	• The opposition **believes** that the policy is ineffective. • مخالفان اعلام کردند که این سیاست ناکار آمد است (صحیح: تغییر گروه فعل در فرایند ذهنی " عقیده داشتن".

The qualitative analysis of the diagnostic phase revealed that learners' translation difficulties were not confined to surface-level lexical inaccuracies but extended to structural, pragmatic, and meta-functional distortions that altered experiential meaning and institutional stance in political discourse. While lexico-semantic errors were the most pervasive, particularly in the rendering of domain-specific terminology and polysemic political expressions, syntactic misconfigurations frequently affected participant roles, temporal relations, and clause organization. More critically, pragmatic shifts and meta-functional deviations, such as agent deletion and process-type reconfiguration, resulted in subtle but consequential changes in agency, modality, and ideological positioning. From a sociocultural and dynamic assessment perspective, these patterns indicate that learners' challenges were developmentally stratified: lower-level lexical issues were generally amenable to implicit mediation, whereas ideational and meta-functional distortions required more explicit, metalinguistic support.

4.1.2. Calibrated Inventory Development

These findings (Table 4.1) substantiate the theoretical premise of the study that translation quality is closely tied to the alignment between mediation type and learners' zone of proximal development, thereby providing an empirical basis for the subsequent design of ZPD-calibrated AI prompts. Therefore, the next step was devoted to the development of an appropriated and calibrated mediational inventory. The mediational inventory was developed within the interactionist model of dynamic assessment, which conceptualizes the zone of proximal development as a dialogic and transformative space integrating assessment and instruction (Poehner & van Compernolle, 2011). Accordingly, mediation was not treated as corrective feedback delivered in a fixed manner, but as contingent, graduated support systematically attuned to learners' emerging needs (Aljaafreh & Lantolf, 1994). Thus, guided by

Farokhipour (2019) and following Poehner and van Compernolle's (2011) framework, the inventory was structured through three successive stages. First, elicitation of learner verbalization was employed to diagnose the source of difficulty. Learners were prompted to explain or justify their translation choices, allowing the mediator to identify conceptual, lexical, or meta-functional breakdowns rather than immediately supplying corrections. Second, a collaborative interactional frame was established in which mediation was targeted to the identified problem and delivered along an implicit-explicit continuum. Support began with minimal prompts (e.g., signaling an inconsistency) and progressed toward more explicit cues (e.g., metalinguistic explanation) only when necessary. Third, a cooperative interactional frame extended mediation beyond immediate error resolution by jointly re-specifying the translation goal or principle involved, thereby fostering transfer and strategic awareness.

In line with established constructivist principles in language learning (Lidz, 1991; Poehner, 2005), the inventory incorporated graduated mediational moves and documented learner reciprocity, understood as the degree of responsiveness to mediation. This ensured that support remained systematic yet flexible, and developmentally calibrated rather than uniformly applied. The resulting inventory (table 4.2) provided a principled framework for operationalizing ZPD-sensitive mediation in the translation tasks and subsequently informed the design of the AI-mediated prompts.

Table 4.2. *ZPD-Calibrated Mediational Moves and Learner Reciprocity in AI-Mediated Translation*

Mediation Level	Type of AI Mediational Move	Illustrative Functions in Political Translation	Expected Learner Reciprocity
Level 1: Implicit (Attention Regulation)	Global prompt; request for reconsideration; eliciting justification	Signals potential inconsistency; prompts re-reading; asks learner to explain lexical, modal, or agency choice	Self-revision; identification of problematic segment; partial reformulation
	Pinpointing segment; narrowing error type; strategic semantic or	Identifies lexical, syntactic, pragmatic, or meta-functional site; prompts reconsideration of process type,	Guided correction; clarification requests; targeted reformulation

Level 2: Focused (Error Localization)	structural hint; binary options	participant roles, or political register	
Level 3: Explicit (Metalinguistic Mediation)	Metalinguistic explanation; SFL-based clarification; register and terminology explanation	Explains agency shifts, passive constructions, modality differences, institutional terminology, or process-type realization	Accurate correction following explanation; verbalized understanding; rule application to adjacent segment
Level 4: Direct (Corrective Modeling)	Provision of correct form; clause reformulation; rule articulation request; pattern review	Supplies accurate translation preserving agency, stance, and discourse function; briefly justifies correction	Acknowledgment; rule restatement; independent application in subsequent task

The mediational inventory presented in table 4.2 was operationalized in the AI-mediated condition through a rule-based prompting sequence calibrated to learners' responsiveness. Drawing on Poehner and van Compernolle's (2011) interactionist model of dynamic assessment, the AI delivered mediation along a graduated implicit-to-explicit continuum, beginning with attention-directing prompts and progressing toward metalinguistic explanation or corrective modeling only when necessary. Learner reciprocity (e.g., self-revision, guided correction, or lack of uptake) determined movement across mediational levels, thereby approximating ZPD-sensitive contingency. The prompts were systematically aligned with recurrent error domains identified in the qualitative analysis (lexico-semantic, syntactic-grammatical, pragmatic, and meta-functional), particularly those involving process-type shifts, participant roles, and political register. This design ensured that AI mediation functioned not merely as automated feedback but as developmentally calibrated instructional support.

4.2. Quantitative Findings

Building upon the qualitative diagnostic analysis and the development of the ZPD-calibrated mediational inventory, the quantitative phase aimed to examine the comparative effectiveness of different mediation modalities on learners' translation performance. Specifically, this phase investigated whether developmentally attuned AI mediation yields differential outcomes compared to normative AI use and conventional teacher-mediated feedback. To this end, a post-test only control group

design was implemented to isolate the impact of mediation type on translation quality while maintaining instructional equivalence across groups. Participants, therefore, were assigned to one of three instructional conditions. The first experimental group (ZPD-tuned AI mediation, N = 10) received adaptive, graduated prompts delivered via ChatGPT, operationalized through the mediational inventory derived from the diagnostic phase. Mediation in this condition followed a systematic implicit-to-explicit continuum and was dynamically adjusted in response to learners' errors, approximating interactionist dynamic intervention principles. The second experimental group (Normative AI mediation, N = 11) interacted with ChatGPT without calibrated scaffolding; prompts were not systematically aligned with learner errors or developmental level, thereby reflecting typical, unguided AI-assisted translation practice. The control group (Teacher-mediated feedback, N = 11) received conventional classroom feedback delivered by the instructor, consistent with established instructional routines rather than ZPD-sensitive calibration.

All groups completed equivalent translation tasks under comparable conditions to ensure procedural consistency. The translation task administered in this phase was deliberately constructed to be dense with problem-prone areas identified in the qualitative diagnostic phase. Specifically, the source text contained lexically ambiguous political terminology, shifts in modality, passive constructions affecting agency distribution, and clauses instantiating varied process types (e.g., material, mental, and relational processes) likely to trigger participant-role distortions. In this way, the task was designed not merely as a performance measure but as a controlled elicitation context targeting previously documented lexico-semantic, syntactic, pragmatic, and meta-functional vulnerabilities. By contrasting calibrated AI mediation, non-calibrated AI assistance, and traditional teacher feedback, this phase quantitatively evaluates whether ZPD-sensitive AI scaffolding produces measurable gains in translation quality beyond normative instructional and technological practices. These findings are reported in the following tables:

Table 4.3. *Comparing Translation Quality across Mediation Conditions (one-way ANOVA)*

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	170.732	2	85.366	33.574	.000
Within Groups	73.736	29	2.543		
Total	244.469	31			

The one-way ANOVA revealed a statistically significant effect of mediation type on translation quality, $F(2, 29) = 33.574$, $p < 001$. The between-group variance (SS = 170.732) substantially exceeded the within-group variance (SS = 73.736), indicating that differences in instructional condition accounted for a considerable proportion of the total variability in scores. These results demonstrate that the type of mediation (ZPD-tuned AI, normative AI, or teacher-mediated feedback) significantly influenced post-test translation performance.

Table 4.4. *Post Hoc Multiple Comparisons of Translation Quality across Mediation Conditions*

(I) Groups	(J) Groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval
					Lower Bound
AI-Tuned Group	Normative-AI	5.15455*	.69672	.000	3.4339
	Teacher Feedback	4.79091*	.69672	.000	3.0703
Normative-AI	AI-Tuned Group	-5.15455*	.69672	.000	-6.8752
	Teacher Feedback	-.36364	.67992	.855	-2.0428
Teacher Feedback	AI-Tuned Group	-4.79091*	.69672	.000	-6.5115
	Normative-AI	.36364	.67992	.855	-1.3155

Besides, post hoc comparisons using Tukey's HSD test indicated that the ZPD-tuned AI group significantly outperformed both the Normative AI group (MD = 5.15, $p < .001$) and the Teacher Feedback group (MD = 4.79, $p < .001$). The 95% confidence intervals for these comparisons did not include zero, confirming the robustness of these differences. In contrast, no statistically significant difference was observed between the Normative AI and Teacher Feedback groups (MD = –0.36, $p = .855$), as reflected by a confidence interval spanning zero. These findings suggest that only the developmentally calibrated AI mediation yielded a statistically meaningful advantage in translation performance.

The findings of the present study strongly support the theoretical claim that the effectiveness of AI-assisted translation depends less on access to technology and more on the quality and developmental attunement of mediation. Across conditions, the ZPD-tuned AI group consistently demonstrated superior translation performance compared with both normative AI use and conventional teacher feedback, while the latter two conditions produced broadly comparable outcomes. This pattern aligns closely with the sociocultural premise that learning is optimized when assistance is calibrated to learners' emerging abilities within the zone of proximal development (Vygotsky, 1978; Poehner, 2008). From a constructivist perspective, mediation is not merely corrective but developmental, integrating assessment and instruction into a single transformative process (Lantolf & Poehner, 2010). The qualitative diagnostic phase of the study revealed layered difficulties across lexico-semantic, syntactic, pragmatic, and meta-functional domains. These error patterns provided the empirical basis for constructing graduated prompts that moved from implicit attention-directing cues to explicit metalinguistic support. Such sequencing directly reflects the regulatory scale proposed by Aljaafreh and Lantolf (1994) and the mediation hierarchy articulated by Lidz (1991), where assistance is progressively intensified only as needed. The superior performance of the calibrated AI condition therefore suggests that prompts function pedagogically when they operate as contingent mediational moves rather than as static instructions. Equally revealing is the absence of meaningful differences between normative AI use and teacher-mediated feedback in quantitative phase. This convergence echoes critiques of non-contingent or interventionist feedback models, in which assistance lacks dialogic negotiation and responsiveness (Poehner, 2008, Farokhipour et al., 2020, Saeedi & Soltani, 2025). Simply supplying corrections, whether delivered by a teacher or by an AI system, does not appear sufficient to promote deeper restructuring of learners' strategic or conceptual understanding. Instead, development seems to depend on reciprocity and contingent interaction that foster learner agency and internalization (Namaziandoust, Farokhipour & Rezai, 2025). In this sense, the present findings suggest that mediation quality, rather than the human or technological source of support, is the decisive factor shaping learning outcomes.

These results of the quantitative phase also extend recent research on prompting behaviors in AI-assisted translation. Zhang et al. (2025) report that students often rely on intuitive and underspecified prompts, with limited contextualization or strategic intent. The present study addresses precisely this gap by operationalizing structured, developmentally informed prompting. The improvement

observed in the ZPD-tuned condition indicates that theoretically grounded, systematic guidance, rather than unguided interaction with generative AI, enhances translation performance. At the same time, the similar outcomes observed in the normative AI and teacher groups corroborate concerns that uncritical reliance on AI outputs may result in superficial engagement and limited learning gains (Zhang et al., 2025). Together, these findings reinforce the need for explicit instruction in how prompts should be designed and sequenced. Besides, the results further converge with model-centered prompt engineering research demonstrating that selective and strategically structured inputs outperform generic or baseline prompting. Mondshine et al. (2025) show that targeted prompt design mitigates translation errors more effectively than undifferentiated approaches. Likewise, Zhang and Fan (2025) demonstrate that prompts tailored to text type and register produce more accurate and contextually appropriate translations. Although these studies focus primarily on optimizing model output, the present findings extend their insights to the learner dimension by showing that selective, calibrated prompting also enhances human performance. In both cases, efficacy derives from strategic design rather than mere technological sophistication.

5. Conclusion

This study examined whether AI prompts calibrated to learners' zones of proximal development enhance translation quality more than unguided AI assistance or teacher feedback. Conceptualizing AI as a mediational artifact, findings showed that developmentally tuned AI mediation significantly outperformed both normative AI use and teacher feedback, with no meaningful difference between the latter two. Results highlight that translation improvement depends on mediation quality, not technology presence alone. Focusing on Iranian undergraduates (N=32) using ChatGPT with political texts, limitations include small sample size, quasi-experimental design, short-term performance only, and no delayed post-tests.

References

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *Modern Language Journal, 78*(4), 465–483. https://doi.org/10.2307/328585.

Farokhipour, S. (2019). *Developing a local model of dynamic assessment for diagnosing problems and promoting English language skills in Iranian context* (Unpublished doctoral dissertation). Chabahar Maritime University.

Farokhipour, S. (2025). Constructivist vs. positive psychology on learning: Developing a positive intervention inventory to mediate speaking, anxiety, and engagement. *Iranian Journal of English for Academic Purposes, 14*(2), 1–18.

Farokhipour, S., Khoshsima, H., Sarani, A., & Ganji, M. (2020). Presenting and investigating the effect of a local model of dynamic assessment in diagnosing and removing learning difficulties of high school students in productive skills. *Journal of Foreign Language Research, 10*(1), 120–134. https://doi.org/10.22059/jflr.2019.273144.593

Farokhipour, S., Padiz, D., & Hekmatshoartabari, B. (2025). Investigating the accuracy of GPT's translation in transferring syntax–semantic interface: A case study of ideational meta-functions in English translation of Quran. *Language Research, 16*(1), 141–166. https://doi.org/10.22059/jolr.2025.391037.666912

Feuerstein, R., Falik, L., & Rand, Y. (2002). *The dynamic assessment of cognitive modifiability*. ICELP Press.

Gao, Y., Wang, R., & Hou, F. (2023). How to design translation prompts for ChatGPT: An empirical study. *arXiv*. https://arxiv.org/abs/2304.02182

Jiao, H., Peng, B., Zong, L., Zhang, X., & Li, X. (2024). Gradable ChatGPT translation evaluation. *arXiv*. https://arxiv.org/abs/2401.09984

Khosravani, M., Farokhipour, S., Padiz, D., & Hekmatshoartabari, B. (2025). A comparison of AI-assisted, AI-revised and human-scaffolded translations in ESP classes. *Translation and Interpreting Research, 2*(5), 45–58. https://doi.org/10.22054/tir.2025.86625.1044

Lantolf, J. P., & Poehner, M. E. (2010). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research, 15*(1), 11–33. https://doi.org/10.1177/1362168810383328

Liang, L., & Lu, S. (2025). The evaluation and reception of the translation quality of three translation modalities in live-streaming contexts: Computer-assisted simultaneous interpreting, machine translation (MT) with human revision and raw MT. *The Translator*. Advance online publication. https://doi.org/10.1080/13556509.2025.2494566

Lidz, C. S. (1991). *Practitioner's guide to dynamic assessment*. Guilford Press.

Mondshine, I., Paz-Argaman, T., & Tsarfaty, R. (2025). Beyond English: The impact of prompt translation strategies across languages and tasks in multilingual LLMs. *Findings of the Association for Computational Linguistics: NAACL 2025*, 1331–1354. https://doi.org/10.18653/v1/2025.findings-naacl.73

Mohsen, M., Althebi, S., & Albahooth, M. (2023). A scientometric study of three decades of machine translation research: Trending issues, hotspot research, and co-citation analysis. *Cogent Arts & Humanities, 10*(1). https://doi.org/10.1080/23311983.2023.2242620

Poehner, M. E. (2005). *Dynamic assessment of oral proficiency among advanced L2 learners of French* (Doctoral dissertation). Pennsylvania State University.

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. Springer.

Poehner, M. E., & van Compernolle, R. A. (2011). Frames of interaction in dynamic assessment: Developmental diagnoses of second language learning. *Assessment in Education: Principles, Policy & Practice, 18*(2), 183–198. https://doi.org/10.1080/0969594X.2011.567116

Saeedi, Z., & Soltani, M. (2025). Developing an AI chatbot for language pragmatics instruction: From algorithms to dynamic assessment in an EFL context. *Computer Assisted Language Learning*, 1–22. https://doi.org/10.1080/09588221.2025.2532014

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, H., & Zhang, J. Z. (2025). Translation practice models in the GenAI era: Technology iteration, business format change, and trend prospect. *Journal of Foreign Language Teaching, 46*(1), 53–58. https://doi.org/10.16362/j.cnki.cn61-1023/h.2025.01.011

Wu, G., & Ding, Z. (2025). An exploration of the linguistic features of translations generated by large language models: An investigation based on average dependency distance and its correlation with translation quality. *Journal of Foreign Languages, 3*, 48–55.

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

Zhang, Y., & Fan, C. (2025). The impact of prompt strategies on specialized text translation by large language models: An empirical study using ChatGPT. In *Proceedings of the 2025 International Conference on Educational Technology and Artificial Intelligence* (pp. 218–222). https://doi.org/10.1145/3766557.3766594