

The Impact of Contextual Prompting on LLM Translation Accuracy in Educational Sciences: An Exploratory Case Study of ChatGPT Pro

Mayssaa Moukatib^{1*} & Rajae Khaloufi²

^{1&2}Abdelmalek Essaâdi, Tangier, Morocco

*Email 1 (Corresponding author): mayssaa.moukatib@etu.uae.ac.ma

Email 2: r.khaloufi@uae.ac.ma

ORCID iD 1 : 0009-0004-9055-4169

ORCID iD 2 : 0009-0001-6624-3019

Received	Accepted	Published online
11/4/2026	17/10/2026	22/10/2026

 : 10.63939/ajts.9ggs0y80

Cite this article as: Moukatib, M., & Khaloufi, R. (2026). The Impact of Contextual Prompting on LLM Translation Accuracy in Educational Sciences: An Exploratory Case Study of ChatGPT Pro. *Arabic Journal for Translation Studies*, 5(17). <https://doi.org/10.63939/ajts.9ggs0y80>

Abstract

Integrating Large Language Models (LLMs) into specialized translation workflows introduces nuanced lexical and register-related challenges. This exploratory case study investigates how progressively constrained prompt engineering influences the translation of domain-specific pedagogical terminology from English into Arabic. Using ChatGPT Pro (GPT-4), an 84-word source excerpt containing four key Task-Based Language Teaching (TBLT) pedagogical terms was translated under three prompting conditions: zero-shot baseline, role-prompting (expert persona), and context-rich prompting (localized to Moroccan secondary education). The resulting Arabic outputs were evaluated against recognized Moroccan pedagogical frameworks and specialized Arabic educational lexicons using an explicit three-tier adequacy rubric. Descriptive analysis revealed marked lexical variations: zero-shot prompting yielded literal semantic calques, whereas context-rich prompting aligned closely with localized curricular terminology. While these exploratory findings demonstrate that contextual constraints assist the model in selecting domain-appropriate registers, the results highlight the necessity of continued human expertise and institutional oversight in specialized machine translation post-editing.

Keywords: Prompt Engineering, Machine Translation, Educational Sciences, ChatGPT Pro, Lexical Accuracy, Exploratory Case Study

1. Introduction

The adoption of Artificial Intelligence (AI), especially Large Language Models like ChatGPT, has radically changed the field of translation and language mediation (Bowker, 2023). These models have achieved high syntactic fluency and coherence in cross-lingual tasks, prompting a shift in translation training from initial drafting to post-editing and evaluating outputs (Koehn, 2020). Nevertheless, while AI shows impressive ability in translating general language, it still faces notable challenges in accurately handling specialized academic and technical content (Castilho et al., 2018).

In Educational Sciences, texts are rich with specific pedagogical terms and jargon. Translating concepts like "Task-Based Learning," "Zone of Proximal Development," or "scaffolding" from English into standard academic Arabic involves more than simple word-for-word translation; it requires cultural adaptation, didactic understanding, and familiarity with educational frameworks in the target language. When LLMs handle these specialized texts without guidance, they often produce literal translations that remove the pedagogical nuance. The main challenge is not the AI's linguistic skills but its ability to grasp the specialized context and convey it accurately.

Previous research has increasingly focused on the relationship between generative AI and translation quality. Recent evaluations of ChatGPT's ability to translate different language pairs aimed to compare its performance with traditional Neural Machine Translation (NMT) tools (Jiao et al., 2023). These studies mainly involved qualitative error analysis of broad academic texts, finding that although LLMs produce more natural syntax, they often struggle with domain-specific terminology. Likewise, foundational research on prompt engineering examined how adding contextual constraints, such as zero-shot versus few-shot learning, impacts AI translation quality (Brown et al., 2020). While this study also aims to assess LLM translation accuracy, it differs mainly in its sample and variables. Instead of analyzing general academic texts, it specifically focuses on complex terminology within educational sciences. Additionally, it introduces progressive contextual constraints, particularly prompt engineering, as a variable to systematically measure effects on lexical accuracy.

This localized challenge in human-machine interaction within translation raises the core research question: How do progressively contextualized prompt engineering strategies influence the accuracy of English-to-Arabic translations of specialized pedagogical terminology by Large Language Models?

This study hypothesizes that by assigning specific expert roles and incorporating localized pedagogical contexts into prompts, literal translation errors can be greatly reduced, leading to highly precise and standardized Arabic educational terminology compared to baseline zero-shot methods. The main aim is to evaluate how varying levels of contextual constraints influence the lexical output of LLMs, highlighting the importance of prompt engineering as a key skill for language experts working with specialized texts.

2. Literature Review

2.1. The Paradigm Shift in Translation Education: From Drafting to Post-Editing

The rapid development of artificial intelligence has significantly disrupted traditional approaches to teaching translation. Historically, translator training focused on linguistic skills, text analysis, and manual drafting as key components of translation ability. However, with the rise of Neural Machine Translation (NMT) and Large Language Models (LLMs), there has been a necessary shift in pedagogy. Moorkens (2018) noted this early on, showing that as machine translation improves to near-human quality in general fields, the human translator's role shifts from creating drafts to post-editing machine-produced texts (PEMT). This change is officially acknowledged by the European Master's in Translation (EMT) network, which now requires graduates to have the skills to critically evaluate, correct, and work with AI-generated outputs (European Commission, 2022). As a result, translator training increasingly employs Task-Based Learning (TBL) methods that reflect contemporary professional settings where human-computer interaction is vital to the translation process.

2.2. LLMs and the Challenge of Specialized English-Arabic Terminology

Although LLMs like ChatGPT exhibit impressive fluency in generating text, their use in specialized domain translation, especially between morphologically different languages like English and Arabic, remains challenging. Studies assessing their translation skills show they perform well in syntactic naturalness but often falter with domain-specific vocabulary. Peng et al. (2023) evaluated ChatGPT's translation capabilities and found that while it matches commercial NMT systems in high-resource languages, it encounters difficulties with technical jargon unless given explicit contextual clues. In fields like Educational Sciences, this issue becomes more pronounced. English educational terms such as "scaffolding" or "Zone of Proximal Development" depend heavily on abstract, domain-specific concepts. When LLMs translate these into Arabic without sufficient context, they tend to produce literal or generic translations (e.g., translating "scaffolding" as construction equipment), which undermines the academic meaning.

2.3. Prompt Engineering as a New Translation Competence

To address lexical inaccuracies, translation research now emphasizes "prompt engineering" or "prompt programming." Instead of viewing the machine as a passive entity that produces fixed translations, users now engage with LLMs interactively. Reynolds and McDonell (2021) demonstrated that LLMs perform best when given contextual constraints, moving beyond "zero-shot" translation, which asks for no context, toward role-playing and detailed instructions. In practice, this involves embedding the "translation brief" directly into the AI prompt. By instructing the AI to act as an expert educational

translator and providing information about the target audience, the machine's statistical output is guided toward the correct specialized register.

2.4. Synthesis and the Current Study

The current literature clearly highlights the necessity for translator education to incorporate AI interaction, emphasizing that LLMs need contextual guidance to accurately interpret specialized terminology. Nevertheless, a significant gap exists in empirical studies that specifically explore how prompt engineering can address lexical ambiguity in English-to-Arabic educational sciences terminology. This study employs zero-shot, role-prompting, and context-rich strategies on a complex, pedagogically focused text to connect translation training with artificial intelligence, aiming to demonstrate that prompt engineering is an essential tool for Arabic translation professionals.

3. Methodology

3.1. Research Design and Variables

Given the targeted nature of this study, an experimental qualitative case-study design was adopted. Rather than seeking statistical generalizability across broad corpora, this research provides an in-depth lexical and register audit of four specialized pedagogical terms embedded within a controlled 84-word didactic passage. The independent variable is the degree of contextual constraint encoded within the prompt (zero-shot, role-prompting, and context-rich prompting), while the dependent variable is the lexical and pragmatic adequacy of the generated Arabic equivalents.

3.2. Sample Selection and Source Material

The 84-word source excerpt (detailed in Appendix 1) was purposefully extracted from peer-reviewed literature on communicative language pedagogy and Task-Based Language Teaching (adapted from Chen & Wang, 2019, p. 122). This passage was selected because it contains an unusually dense cluster of constructivist and socio-cultural didactic terms within a compact syntactic structure.

3.3. Instruments and Data Collection Procedures

Data generation was conducted on 14 February 2026 using the ChatGPT Pro web interface (OpenAI, February 2026 release, running the GPT-4 architecture). Default generation parameters were maintained (temperature estimated at ≈ 0.7 , Top-p = 1.0, default system presence penalty = 0). Each prompt condition was executed within a fresh, isolated chat session with web-browsing capabilities disabled to ensure that generation relied solely on parametric pre-trained memory and prompt-specified constraints, eliminating cross-session contextual contamination.

Methodological Note on Stochasticity: Because Large Language Models operate on probabilistic token distributions, their outputs exhibit inherent stochasticity across distinct

generations. To assess lexical stability, three consecutive runs were executed per prompt condition. The outputs presented in Table 3 represent the consensus generation observed across these iterations.

The data collection process consisted of inputting the identical 84-word English source text into the LLM across the three specified prompt levels, detailed in Table 1:

Table 1: Progression of Experimental Prompting Constraints

Prompt Level	Category	Verbatim Prompt Instruction	Theoretical Rationale / Added Constraint
Prompt 1	Baseline Zero-Shot	"Translate this text into Arabic: [source text]"	Pure bilingual decoding task without domain cues, persona, or target audience constraints.
Prompt 2	Role-Prompting	"Act as an expert translator in educational sciences. Translate this text into standard academic Arabic: [source text]"	Imposes a professional persona and restricts the target register to general academic discourse, activating domain-specific sub-networks.
Prompt 3	Context-Rich	"Translate this text into Arabic for Moroccan high school educators. Ensure educational terminology aligns with standard pedagogical frameworks: [source text]"	Imposes localized geographic and institutional constraints (<i>skopos</i>), defining both the recipient group and ministerial curricula.

3.4. Data Analysis Framework

After data collection, the three Arabic translations were compared through a qualitative analysis. The texts were arranged side-by-side in a table to focus on four key pedagogical terms: ZPD, scaffolding, explicit instruction, and cognitive work. A descriptive analytical method was used to summarize and assess the data. To ensure empirical rigor and address the subjectivity of translation evaluation, accuracy was measured using an explicit three-tier adequacy rubric. This rubric systematically cross-references the machine outputs against established Arabic educational dictionaries and official Moroccan didactic frameworks (e.g., CRMEF modules):

Table 2: Three-Tier Lexical Adequacy Rubric

Score	Category	Operational Definition
0	Inadequate / Calque	Literal semantic transfer that violates pedagogical conventions, administrative mistranslation, or inability to translate technical acronyms.
1	Semantically Viable / Generic	Communicatively intelligible paraphrase or general academic equivalent, but lacking specific pedagogical standardization.
2	Standardized Pedagogical Equivalent	Exact alignment with codified Moroccan educational frameworks (CRMEF) and specialized Arabic pedagogical lexicons.

4. Results and their discussion

A quantitative analysis shows a clear link between the use of contextual prompts and the accuracy in terminology. Among the four targeted educational sciences terms, Prompt 1 (Zero-Shot) had a 0% success rate, failing to produce any standardized pedagogical equivalents. Prompt 2 (Role-Prompting) reached 75%, correctly resolving most semantic

ambiguities but not always ensuring exact localization. Prompt 3 (Context-Rich) achieved full success with 100%, accurately translating all four terms into standardized Moroccan pedagogical Arabic.

This study primarily aimed to examine how different levels of contextual constraint (prompt engineering) influence the lexical accuracy of LLMs when translating specialized educational sciences terminology from English to Arabic. To do this, an 84-word source text on Task-Based Language Teaching (TBLT) was processed with ChatGPT Pro using three different prompting strategies. The results are shown below, followed by a detailed qualitative analysis of the variations in terminology.

4.1. Comparative Translation Data

Table 3 shows the complete Arabic translations produced by the LLM across the three experimental conditions. The original English source text used for this generation is included in Appendix 1.

Table 3. Complete Arabic Machine Translation Outputs Across the Three Experimental Conditions

Source Segment (Chen & Wang, 2019) [source: 9, 12]	Prompt 1: Zero-Shot Baseline [source: 9, 12]	Prompt 2: Role-Prompting (Expert Persona) [source: 9, 12]	Prompt 3: Context-Rich (Localized Moroccan Framework) [source: 9, 12]
"Tasks must be structured in a way that they pose an appropriate challenge by requiring learners to perform functions and use language that enable them to dynamically construct ZPDs." [source: 9, 12]	يجب هيكلة المهام بطريقة تشكل تحدياً مناسباً من خلال مطالبة المتعلمين بأداء وظائف واستخدام لغة تمكنهم من بناء ZPDs بشكل ديناميكي.	يجب أن تُبنى المهام التعليمية بحيث تشكل تحدياً ملائماً يتطلب من المتعلمين أداء وظائف واستخدام لغة تتيح لهم البناء الديناميكي لمنطقة النمو الوشيك (ZPDs).	يتعين تنظيم المهام التعليمية بطريقة تطرح تحدياً ملائماً، يستلزم من المتعلمين إنجاز أنشطة واستعمال لغة تمكنهم من البناء الدينامي لمناطق النمو التقاربي.
"Thus, in this process, explicit instruction could be the most appropriate approach to achieve this objective as it helps students scaffolding certain knowledge based on what they already learned." [source: 9, 12]	وبالتالي، في هذه العملية، يمكن أن تكون التعليمات الصريحة هي النهج الأكثر ملاءمة لتحقيق هذا الهدف لأنها تساعد الطلاب في سقالات معرفة معينة بناءً على ما تعلموه بالفعل.	ومن ثم، يُعد التعليم المباشر (الصريح) المقاربة الأكثر ملاءمة لتحقيق هذا الهدف، حيث يساعد المتعلمين في البناء التدريجي لمعارف معينة انطلاقاً من مكتسباتهم السابقة.	ومن هذا المنطلق، يمثل التدريس الصريح المدخل الأكثر ملاءمة لبلوغ هذه الغاية، لكونه يزود المتعلمين بدعامات التعلم اللازمة لبناء المعارف الجديدة وفق مكتسباتهم القبلية.
"[...] During the sequences, the teachers understand learners' thinking, provide scaffolding and stimulate them in cognitive work, gradually helping them comprehend it." [source: 9, 12]	[...] خلال التسلسلات، يفهم المعلمون تفكير المتعلمين، ويقدمون سقالات ويحفزونهم في عمل معرفي، مما يساعدهم تدريجياً على فهمه.	[...] وأثناء المقاطع التعليمية، يستوعب المدرسون طريقة تفكير المتعلمين، ويقدمون الدعم اللازم ويحفزونهم على بذل جهد ذهني، لمساعدتهم تدريجياً على استيعاب المحتوى.	[...] وخلال المقاطع الديدكتيكية، يرصد الأساتذة سيورات التفكير لدى المتعلمين، موفرين لهم الدعم المرحلي ومحفزين إياهم على الانخراط في السيرورة المعرفية لتملك المعارف تدريجياً.
	0/4 Adequate (Score: 0/8)	3/4 Adequate (Score: 5/8)	4/4 Adequate (Score: 8/8)
Evaluative Summary	Contains untranslated acronyms, semantic calques (سقالات), and literal register violations [source: 9, 12].	Semantically accurate general Arabic; avoids calques via paraphrase, but misses institutional codification [source: 9, 12].	Full alignment with codified Moroccan educational didactics (التدريس الصريح، الدعم المرحلي، السيرورة المعرفية). [source: 9, 12].

Note: Evaluated against Moroccan CRMEF didactic standards and specialized pedagogical lexicons (Al-Sharifi, 2000; CRMEF, n.d.-a; Ministry of National Education, 2024a, 2024b). Prepared by the authors (2026).

4.2. Lexical and Semantic Analysis

A detailed descriptive examination of the generated texts highlights notable semantic and register shifts across the four targeted pedagogical terms. The analysis below outlines how the model outputs resolved these specific concepts under varying prompt constraints.

4.2.1. ZPDs (Zone of Proximal Development)

This core Vygotskian concept illustrates the constraints of unrestricted generative translation.

- **Prompt 1** exhibited an inability to resolve the Vygotskian acronym, leaving "ZPDs" untranslated within the Arabic text, creating an immediate cognitive barrier for the target reader.
- **Prompt 2** correctly unpacked the acronym into standard academic Arabic, rendering it as "منطقة النمو الوشيك", which is the conventional Middle-Eastern translation (Hamdan, 2007; UNESCO-IBE, 2013).
- **Prompt 3** generated the highly specific localized term "منطقة النمو التقاربي". This directly mirrors the Francophone didactic heritage (*Zone de proche développement*) codified in Moroccan teacher training modules (Regional Centers for Education and Training Professions [CRMEF], n.d.-b).

4.2.2. Scaffolding

The metaphor of instructional "scaffolding" revealed significant literal transfer vulnerabilities in the baseline model.

- **Prompt 1** produced an unacceptable literal calque ("سقالات"), denoting physical construction site scaffolding (Almahasees, 2021). Within an educational context, this disrupts the academic integrity of the text.
- **Prompt 2** generated explanatory circumlocutions ("بناء المعرفة التدريجي" and "الدعم"), capturing the communicative intent but failing to achieve technical brevity.
- **Prompt 3** successfully generated the standardized pedagogical terminology utilized in regional teacher training centers, producing "دعامات التعلم" (learning supports) and "الدعم المرحلي" (phased instructional support). This aligns precisely with Moroccan didactic frameworks (Ministry of National Education, Preschool and Sports, 2024b; CRMEF, n.d.-a).

4.2.3. Explicit Instruction

This term requires distinguishing between general directives and formal teaching methodologies.

- **Prompt 1** yielded a translation that conflated the pedagogical concept with administrative directives, mistranslating it as "تعليمات صريحة" (explicit orders).
- **Prompts 2 and 3** correctly identified the didactic methodology. Prompt 3 specifically employed "التدريس الصريح", which is the exact pedagogical terminology mandated in Moroccan educational reforms and teacher qualification programs for targeted skill acquisition (Ministry of National Education, Preschool and Sports, 2024a; Al-Sharifi, 2000).

4.2.4. Cognitive Work

While all three outputs conveyed the overall meaning, the level of stylistic and domain-specific refinement varied based on the constraints.

- **Prompt 1** produced a rigid, word-for-word translation ("عمل معرفي").
- **Prompt 2** offered a colloquial paraphrase ("جهد ذهني" - mental effort).
- **Prompt 3** elevated the register to "السيرورة المعرفية" (the cognitive process), matching the constructivist terminology standardized in Moroccan educational psychology and evaluation curricula (Ben Chtioui, n.d.; Sellami et al., 2019).

4.3. Discussion

The empirical data above provides preliminary descriptive support for the hypothesis: using context-rich prompt engineering notably decreases literal translation errors and greatly enhances terminological accuracy of LLMs in specialized fields.

When compared to existing research, these findings provide a practical approach to addressing the domain-specific weaknesses identified in earlier LLM benchmarks (e.g., Jiao et al., 2023). Zero-shot prompting (Prompt 1) prompts the AI to depend on the most common token associations from its extensive training data, which often results in literal translations like "سقالات" for scaffolding. In contrast, framing the prompt as a formalized "translation brief" (Prompt 3) helps the human operator guide the LLM to generate outputs in the correct academic style.

These findings indicate that the role of the modern translator is shifting from manual drafting to designing AI-context frameworks. By dedicating time to develop strong, domain-specific prompts beforehand, language professionals can avoid common lexical errors in machine-translated specialized texts, making post-editing more efficient and maintaining high-quality, localized academic communication.

4.4. Methodological Reflexivity: Skopos Specification vs. Target Leakage

A critical methodological consideration in prompt engineering research is the boundary between providing contextual parameters (*skopos*) and indirectly 'leaking' the target solution into the prompt. In this experiment, Prompt 3 instructed the model to align with 'standard pedagogical frameworks for Moroccan high school educators' without providing any specific Arabic lexical items or glossaries. Thus, the prompt did not furnish the model with the answers; rather, it functioned as an authentic translation brief (Nord, 1997). By

specifying the communicative purpose, target readership, and institutional jurisdiction, the prompt directed the LLM's latent probabilistic weights away from pan-Arab generic internet corpora and toward the localized Franco-Arabic didactic sub-register characteristic of Moroccan education. Rather than confounding the results, this outcome illustrates the core mechanism of contextual prompting: generative LLMs possess dormant domain-specific competencies that remain unactivated under zero-shot conditions, requiring explicit situational parameters to emerge.

5. Conclusion

This study aimed to address a key question in modern translation education: how does using progressively more contextualized prompt engineering influence the accuracy of English-to-Arabic translation of specialized pedagogical terms by Large Language Models? Results clearly show that relying on unconstrained, zero-shot AI prompts for technical texts causes significant lexical errors and literal translations that misrepresent academic meaning. Conversely, guiding the LLM with context-rich prompts, such as those assigning expert personas and specifying regional pedagogical frameworks, helps the model avoid literal pitfalls and produce precise, standardized Arabic educational terminology. Essentially, the prompt acts as a digital translation brief, effectively shaping the necessary semantic and pragmatic tone.

This study suggests a change in translation education by emphasizing 'prompt engineering' as a key skill. Translation programs and regional centers should formally include this in their curricula. Rather than viewing machine translation engines as passive tools that produce rough drafts requiring extensive post-editing, trainee translators should learn to actively configure the AI's contextual settings before generating text. This approach can save time and help maintain terminological consistency.

Although these promising results are evident, the research recognizes certain theoretical and practical constraints. Methodologically, it employs a targeted qualitative case-study approach with a small, purposive textual sample consisting of an 84-word excerpt. Additionally, the experiment was limited to only one generative model, ChatGPT Pro. Therefore, while the lexical analysis offers valuable qualitative insights into AI behavior, the results are not statistically generalizable to other text types or different machine translation systems.

Looking ahead, these limitations present multiple opportunities for future research. Subsequent studies could enhance this framework by integrating mixed-methods approaches, utilizing prompt engineering techniques on large collections of educational texts, and performing quantitative error analysis with statistical tools. Furthermore, researchers are encouraged to undertake comparative studies that evaluate the performance of ChatGPT's prompt-engineered outputs against other neural translation engines like DeepL across various specialized fields, thereby contributing to a better understanding of the evolving dynamics in human-computer interaction within professional translation.

Acknowledgements

The authors wish to acknowledge the valuable contributions of scholars and researchers whose work on translation studies, educational terminology, and artificial intelligence has informed the theoretical and methodological foundations of this study. Particular recognition is given to the Moroccan pedagogical and curricular frameworks that provided essential reference points for evaluating the contextual and terminological accuracy of the machine-generated translations.

Ethical approval

This study did not involve human participants, interviews, surveys, or the collection of personally identifiable information. The research was based exclusively on textual analysis and the experimental evaluation of AI-generated translations of a published educational text. Accordingly, formal ethical approval was not required for the research procedures described in this study.

Author contributions

Mayssaa MOUKATIB: Conceptualization, research design, methodology, investigation, experimental data collection, formal analysis, and writing of the original manuscript.

Rajae KHALOUFI: Methodological guidance, supervision, validation of the analytical framework, interpretation of findings, and critical review and editing of the manuscript.

Both authors contributed to the intellectual development of the study, reviewed the manuscript, and approved the final version for publication.

Disclosure statement

The authors declare that they have no known financial, professional, or personal conflicts of interest that could have influenced the research, analysis, or conclusions presented in this article. ChatGPT Pro was employed as an experimental tool to investigate the impact of contextual prompting on the accuracy of English-to-Arabic translations of specialized educational terminology. The authors maintained full responsibility for the research design, evaluation of AI-generated outputs, interpretation of findings, and scientific integrity of the study.

Funding

This research received no external funding. The study was conducted independently, without financial support from public, commercial, or non-profit funding agencies.

About the authors

Mayssaa MOUKATIB is a PhD researcher specializing in Translation Studies and Linguistics at the King Fahd School of Translation, Abdelmalek Essaâdi University,

Morocco. She is also a secondary-school English teacher. She holds a Bachelor's degree in English Studies, specializing in Linguistics, and a Master's degree in Translation Studies and Linguistics. Her academic interests include artificial intelligence in translation, human-machine translation, machine translation post-editing, translator education, and the translation of specialized educational terminology. Her research particularly explores the integration of emerging AI technologies into translation practices and pedagogical frameworks.

Rajae KHALOUFI is an Associate Professor and translator at the King Fahd School of Translation, Abdelmalek Essaâdi University, Morocco. She holds a PhD in Linguistics, Communication and Translation. Her academic and research interests encompass Translation Studies, Literary Translation, Performance Studies, and intercultural communication. Her scholarly contributions examine the intersections of translation, cultural representation, and artistic expression, with particular attention to the role of translation in facilitating intercultural understanding and knowledge exchange.

ORCID

Mayssaa MOUKATIB  <https://orcid.org/0009-0004-9055-4169>

Rajae KHALOUFI  <https://orcid.org/0009-0001-6624-3019>

Data Availability Statement

The data supporting the findings of this study are presented within the article. The original English source excerpt is provided in Appendix 1, while the experimental prompting strategies, lexical adequacy assessment criteria, and comparative Arabic translation outputs are documented in Tables 1, 2, and 3, respectively. The reported translation outputs represent the consensus generations obtained across three experimental runs for each prompting condition. The article does not specify a separate public repository containing the complete individual-run outputs.

References

- Almahasees, Z. (2021). *Analysing English-Arabic machine translation: Google Translate, Microsoft Translator and Sakhr* (1st ed.). Routledge.
<https://doi.org/10.4324/9781003191018>
- Al-Sharifi, S. A. (2000). *Dictionary of educational sciences terms: Arabic - English / English - Arabic*. Obeikan Publishing.
- Ben Chtioui, M. (2014). *Educational assessment*. Regional Centers for Education and Training Professions (CRMEF).

- Bowker, L. (2023). Translation technology and the changing role of the translator. In F. Zanettin & C. Rundle (Eds.), *The Routledge handbook of translation and technology* (pp. 411-426). Routledge. <https://doi.org/10.4324/9781315311258-26>
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877-1901. <https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf>
- Castilho, S., Doherty, S., Gaspari, F., & Moorkens, J. (2018). Approaches to human and machine translation quality assessment. In J. Moorkens, S. Castilho, F. Gaspari, & S. Doherty (Eds.), *Translation quality assessment: From principles to practice* (pp. 9-38). Springer. https://doi.org/10.1007/978-3-319-91241-7_2
- Chen, S., & Wang, J. (2019). Effects of task-based language teaching (TBLT) approach and language assessment on students' competences in intensive reading course. *English Language Teaching*, 12(3), 119-138. <https://doi.org/10.5539/elt.v12n3p119>
- European Commission. (2022). *European Master's in Translation competence framework 2022*. Directorate-General for Translation. https://commission.europa.eu/system/files/2022-11/emt_competence_fw_k_2022_en.pdf
- Hamdan, M. (2007). *Dictionary of education and teaching terminology: Arabic - English*. Dar Knoz Al-Maarifa.
- Jiao, W., Wang, W., Huang, J., Wang, X., & Tu, Z. (2023). Is ChatGPT a good translator? A preliminary study. *arXiv*. <https://doi.org/10.48550/arXiv.2301.08745>
- Koehn, P. (2020). *Neural machine translation*. Cambridge University Press. <https://doi.org/10.1017/9781108608480>
- Ministry of National Education, Preschool and Sports. (2024a). *ANG COLL.046-24: Qualifying training for middle school English teachers* [Official directive]. Kingdom of Morocco.
- Ministry of National Education, Preschool and Sports. (2024b). *Ministerial decree concerning the organization of qualifying training for primary education teachers* (Decree No. 22 November 2024). Kingdom of Morocco.
- Moorkens, J. (2018). What to expect from neural machine translation: A practical in-class translation evaluation exercise. *The Interpreter and Translator Trainer*, 12(4), 375-387. <https://doi.org/10.1080/1750399X.2018.1501639>
- Nord, C. (1997). *Translating as a purposeful activity: Functionalist approaches explained*. Routledge.
- OpenAI. (2024). *GPT-4 technical report*. arXiv. <https://doi.org/10.48550/arXiv.2303.08774>

- Peng, K., Ding, L., Zhong, Q., Shen, L., Liu, X., Zhang, M., Ouyang, Y., & Tao, D. (2023). *Towards making the most of ChatGPT for machine translation*. arXiv. <https://doi.org/10.48550/arXiv.2303.13780>
- Regional Centers for Education and Training Professions (CRMEF). (2022). *Module descriptor card: Planning*. Ministry of National Education, Preschool and Sports.
- Regional Centers for Education and Training Professions (CRMEF). (2022). [Management module: Teaching and learning situations]. Ministry of National Education, Preschool and Sports.
- Reynolds, L., & McDonell, K. (2021). Prompt programming for large language models: Beyond the few-shot paradigm. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 1-7. <https://doi.org/10.1145/3411763.3451760>
- Sellami, A. L., Sawalhi, R., Romanowski, M. H., & Amatullah, T. (2019). Definitions of educational leadership: Arab educators' perspectives. *International Journal of Leadership in Education*, 25(5), 767-786. <https://doi.org/10.1080/13603124.2019.1690701>
- UNESCO International Bureau of Education. (2013). *Glossary of curriculum terminology*. UNESCO-IBE. <https://2h.ae/QuNQe>

Appendix 1

Tasks must be structured in a way that they pose an appropriate challenge by requiring learners to perform functions and use language that enable them to dynamically construct ZPDs. Thus, in this process, explicit instruction could be the most appropriate approach to achieve this objective as it helps students scaffolding certain knowledge based on what they already learned. [...] During the sequences, the teachers understand learners' thinking, provide scaffolding and stimulate them in cognitive work, gradually helping them comprehend it.