

# Artificial Intelligence or Peer Discussion? Investigating Learning Outcomes in Medical Translation Tasks

Khaled S. Aldossary

Department of English Language, College of Arts, King Faisal University, Al Ahsa, Saudi Arabia

**Abstract**—With the growing integration of artificial intelligence (AI) in translation education, understanding its influence on learning outcomes has become increasingly important. This mixed-methods study examined the effects of different types of translation tasks on participants' ability to learn and remember terminology. Participants consisted of 48 undergraduate students taking a medical translation course. Half collaborated in traditional group tasks, while the rest employed the large language model (LLM) tools DeepL and ChatGPT. Progress was measured quantitatively by having each participant take pre-, immediate post-, and delayed post-tests. In addition, semi-structured interviews were conducted with six of the participants to gather qualitative data. Participants showed significant improvement on the initial post-test regardless of the group they were assigned to. On the delayed post-test, however, the participants doing group work maintained this improvement, while the participants using LLM tools suffered a significant drop in test scores. These results indicated that either strategy could work in the short run, but using LLM tools was not an effective long-term strategy. This aligned with the interviews, in which group work participants highlighted the value of discussing tasks and supporting each other, in contrast to those using LLM tools, who pointed out an absence of peer interaction, despite how convenient they found the tools to be. These findings suggest that group work in a medical translation course could help students acquire and maintain new vocabulary, although using LLM tools could have short-term benefits as well.

**Index Terms**—collaborative learning, medical translation, ChatGPT, DeepL, vocabulary retention

## I. INTRODUCTION

The demand for modern translation and interpretation is on the rise as language barriers continue to vanish. In the medical field, this is more so in a global setting where effective practice requires efficient communication. Previously, the medical profession relied on human translation or strict adherence to medical services where preferred languages are offered. This has since changed with the introduction of AI-Translation tools (Ratna, 2019). Medical translation is a specialized field and it requires the translation of medical terminology from one language to another. Terminology retention is an important aspect in this field, and a variety of research has been carried out in this regard (Saeed et al., 2025; Csaba et al., 2025).

AI has opened new avenues in language learning and professional translation, with medical translation emerging as one of its most promising applications. Precise language interpretation in the clinical domain is fundamental in preventing misunderstanding, preventable errors, adverse effects and harm to patients. Hospitals make use of the services of translators (Genovese et al., 2024). More efficiently, clinical settings provide the services of AI translators. AI in the medical field performs the activities that usually need human intelligence. It addresses the deficiency in human availability and provides a more cost-effective solution. The research on AI in medical translation is novel. Its use is explored on devices such as computers, tablets, and smartphones. This allows both parties (patients and health workers) to communicate despite language barriers.

Terminology retention was strictly dependent on crude methods among students: flash cards, breaking of root words, prefixes and suffixes, mnemonics and visualization, consistent practice with real life medical charts (Alshehri, 2024; Brahler & Walker, 2008). More recently, AI has opened new avenues. Both medical terminology retention and translation are intertwined. Translating complex terms requires deconstructing terminology into components – prefixes, roots and suffixes.

Medical terminology is considered one of the facets of medical translation. Terminology constitutes one of the most challenging facets of medical translation (Karwacka, 2015). A variety of strategies are used by medical translators. This is similar to translators who deal with other disciplines. Montalt-Resurrecció and Shuttleworth (2021) explored the intricate relationship between translation and medical terminology. They concluded that it is encouraged to extend simple linguistics transfers to encompass social, cultural and practical aspects of communication. This interdisciplinary dialogue cuts across sub-disciplines. Standardized medical terminologies are used to ensure accurate and consistent communication of information and to facilitate data exchange. Currently, many terminologies are only available in English, which hinders international research and automated processing of medical data. While these are useful end-points, collaborative learning

has been in existence as long as terminology and translation learning has, and cannot be abandoned. Learning medical terminology and translation in medical classroom is dependent on Instructor feedback and collaborative learning. While the use of AI encourages strict individual usage with software, collaborative learning has been in existence at inception and is now being incorporated (Mossop et al., 2019, pp. 136–157).

Collaborative tasks have been shown to facilitate pragmatic knowledge while helping learners engage in translation at a deeper cognitive level, thereby helping them remember terms (Kargar & Ahmadi, 2021). Furthermore, students following this approach have shown stronger progress than peers who translated medical texts on their own (Bayraktar-Özer & Hastürkoğlu, 2020). This approach can help learners become more self-confident, motivated, and proficient in collaborative technology (e.g., Olvera-Lobo et al., 2009; Sánchez Ramos, 2019). In contrast, LLMs have the potential to make translators faster, more autonomous, and more efficient. In a study by Kruk and Kałużna (2024), students who employed LLM tools claimed to feel more confident and accurate, especially when they received structured feedback. In another study, Özyurt (2024) found that English instruction employing LLMs facilitated students' acquisition and use of medical terms.

Despite the promise of LLMs, there is a concern that they could result in students having fewer chances to collaborate and achieve deeper cognitive engagement with translation tasks. To address such concerns, this study sought to compare how collaborative and LLM-assisted learning strategies influenced the long-term retention of medical terminology among translation students. Another goal was to see how students perceived these approaches. To achieve these objectives, the study sought to answer the following research questions:

1. Do collaborative translation tasks help students retain medical terminology more than non-collaborative translation tasks facilitated by LLMs?
2. How do students perceive these two types of tasks?

## II. LITERATURE REVIEW

### A. Collaborative Learning

Based on sociocultural theory (Vygotsky, 1978), a collaborative approach frames language learning as being socially constructed. That is, learners work together to actively build knowledge, benefiting from scaffolding and the zone of proximal development, in which they progress with the support of more knowledgeable individuals, who could be teachers or fellow learners. Such an approach can encourage learners to engage higher-order thinking skills (Curtis, 1999; Gokhale, 1995).

Furthermore, collaborative activities have been shown to promote language students' communication skills, creativity, and autonomy (Chyzhykova, 2021). This approach can lead to more student-centered learning, with students actively pursuing knowledge instead of simply absorbing it from lectures. The very act of social interaction—which involves students giving each other feedback, solving problems, and negotiating meaning—can help students better understand what they learn and retain language gains.

In one study, Thai university students exposed to online collaborative translation tasks significantly improved their English ability and reported holding a very positive opinion of this approach (Kitjaroonchai et al., 2018). In another study, collaborative tasks appeared to help students be more engaged in using the second language and feel more confident in their ability (Aldossary, 2025).

However, this approach is not without obstacles and, as Iqbal and Sheeraz (2021) noted, its success requires that tasks are designed clearly, students work together effectively in an inclusive environment, and teachers provide necessary support and guidance. At the same time, this approach can be less effective if students do not participate equally in collaborative tasks, the tasks are dominated by some students, or participants fail to prepare properly.

### B. Automation Tools in Medical Translation Studies

Digital tools have allowed more elements of the translation process to be automated, changing the nature of translator education and training. Glossaries and dictionaries, for instance, have given way to digital term bases and translation memories. Furthermore, machine translation allows students to automatically generate suggested translations. Xu's (2024) findings indicated that LLMs simplified the translation process, while Yuxiu (2024) found that neural machine translation was associated with significantly higher test scores and satisfaction among students. Such tools showed higher accuracy than regular machine translation and were perceived by teachers and students as being easier to use.

Recent advances have enabled LLMs such as ChatGPT to translate text with increasing coherence and fluency. Such tools have led translation students to focus more on post-editing, which requires critical thinking (Chen, 2024). On the other hand, there is concern that users will become too reliant on these tools, leading to a loss of valuable translation skills, not to mention potential errors in LLM-generated medical content. LLMs can make translating medical texts faster and more consistent, but due to the highly specialized nature of such documents, relying too heavily on automation can lead to faulty translations, especially when cultural nuance is involved (Boulanger, 2024).

In a systematic review of machine translation tools (e.g., Google Translate) used by medical personnel, Genovese et al. (2024) showed they tended to be more accurate when going from English to another language and less accurate otherwise. Patients tended to be more satisfied with the tools than medical staff, pointing to such tools becoming increasingly accepted in medical settings. The above trend has led to calls for more education on these tools. Chen (2024)

and Xu (2024), for instance, stated that translation education should find a balance between new and traditional skills in order to prepare students to use these tools appropriately. Thus, these tools can help make the translation process more efficient, but educators and translators need to take care in how they use them, as they are not a drop-in replacement for human oversight.

Collaborative and LLM-assisted translation tasks have both shown promise in previous studies. Collaborative learning has been shown to encourage students to engage with the material and negotiate meaning with peers. LLMs can streamline the translation process, but it remains unclear to what extent, if any, they can benefit students cognitively and socially the way collaborative learning can. It is thus uncertain how pedagogically effective these two strategies are in relation to each other, particularly in an area like medical translation that demands stricter quality control. To address that gap, the present study compared how each strategy affected students' medical terminology retention and elicited their feedback.

### C. *Medical Translation Apps: Google Translate (GT), ChatGPT and DeepL*

Effective translation is required due to the need to increase communication between patients and healthcare professionals. Although ChatGPT and Google Translate handle the majority of the work, they fall short when it comes to complex medical terms. In order to overcome linguistic obstacles, ad hoc machine translation must be used. Language-minority groups now have access to cutting-edge medical services because of this. Having timely access to motivating research ensures that they receive comprehensive care and the information they require in an efficient manner (Noll et al., 2025; Rao et al., 2024).

No translation tool has been specifically built for medical usage or application. The majority of the population uses Google Translate and ChatGPT. While these apps have shown to be effective thus far, creating an AI-driven translation tool based on medical language models could provide far more accurate and trustworthy translation. One of the apps that have aimed to bridge this gap is the DeepL. Although, it is not solely used in medical applications, it has proven to be more useful. It supports many professions: healthcare, pharma and life sciences, legal, governments, financial, amongst others. Like GT and ChatGPT, it supports a train of languages: German, French, Spanish, Chinese, Arabic, and many others (DeepL, 2025; OpenAI, 2025).

The Translation versions are available in DeepL Voice, which ensures real-time voice collaboration to ensure effective communication; and DeepL Write, which offers translation with AI-powered edits to improve writing, spelling, grammar and punctuation. DeepL, along with other platforms that are being newly invented for medical translation centralize on creating benefits by blending cutting-edge research and development with language expertise, essentially breaking language barriers between healthcare professional and patients. This benefits not just medical applications, but also the medical research sector, since it makes previously inaccessible data available due to the absence of language obstacles projected by language-minorities. (Kornbluth et al., 2021; Patil & Davies, 2014).

## III. METHODOLOGY

### A. *Participants*

Participants consisted of 56 undergraduate students taking a medical translation course at a Saudi public university. Each participant was assigned randomly to one of two groups: one that engaged in collaborative translation tasks ( $n = 27$ ) and one that engaged in LLM-assisted tasks ( $n = 29$ ). Students who did not attend all instructional sessions or completed all tests were excluded from data analysis, yielding a final sample of 48, evenly divided between the two groups. In addition, the researcher purposively selected three participants from each group to interview.

### B. *Data Collection*

This study employed two instruments to collect data. The first consisted of translation proficiency tests (pre-, immediate post-, and delayed post-tests). Test items included fill-in-the-blank, multiple choice, and short translations. The purpose was to quantitatively gauge students' acquisition and retention of medical terminology after the treatment (i.e., engaging in either collaborative or LLM-assisted tasks). The second instrument was a semi-structured interview given to six participants to explore how they perceived the treatment.

Before the treatment, the participants were given a pre-test on medical terminology. After that, the treatment was conducted over the course of six weeks (Weeks 1–6). During that time, participants in the collaborative group worked in groups of four on collaborative translation tasks. They were allowed to use online dictionaries only such as Stedman's Medical Dictionary, Merriam-Webster. During the same period, the LLM-assisted group did equivalent tasks on their own with the help of tools such as Google Translate, DeepL, and ChatGPT. The tasks came from students' textbooks and materials for the class. At the end of the treatment phase in Week 6, participants took an immediate post-test to assess any changes in terminology retention. A month after the treatment (Week 10), participants took a delayed post-test to see if there were longer-term changes to terminology retention. The following week (Week 11), the researcher conducted semi-structured interviews with six participants regarding the treatment.

### C. *Data Analysis*

The quantitative data, taken from test scores, underwent analysis to determine any within-group or between-group effects from the treatment on students' ability to retain medical terminology. First, the researcher computed means and

standard deviations for how participants did across the tests. To find differences within groups, the researcher carried out paired-sample *t*-tests for pre-test and immediate post-test results as well as immediate and delayed post-test results. These comparisons were done to find any significant changes in vocabulary knowledge and retention over the course of the study.

Furthermore, the two groups were compared with each other in each of the three testing phases through independent-sample *t*-tests. This made it possible to identify whether there were significant differences between groups. The pre-test comparison showed the groups were equivalent before the treatment, the immediate post-test compared short-term learning gains, and the delayed post-test showed any long-term gains. To make sense of the qualitative interview data, the researcher employed thematic analysis via inductive coding. This process involved iterative reading and coding to find major themes in the treatment's perceived benefits and challenges. The qualitative and quantitative data were then combined for a more complete exploration of the findings.

#### IV. RESULTS

##### A. Quantitative Results

###### (a). Descriptive Statistics

Table 1 shows the descriptive statistics for each group's test scores. Each group showed improvement at the end of the treatment. A month later, however, the collaborative group was able to keep these gains, while the LLM group showed a drop in performance.

TABLE 1  
TEST SCORE MEANS (OUT OF 48) AND STANDARD DEVIATIONS BY GROUP

Group	Pre-Test ( $M \pm SD$ )	Post-Test ( $M \pm SD$ )	Delayed Post-Test ( $M \pm SD$ )
Collaborative	29.40 $\pm$ 2.38	36.04 $\pm$ 2.67	36.22 $\pm$ 2.67
LLM	29.80 $\pm$ 2.55	36.62 $\pm$ 2.32	34.20 $\pm$ 3.04

###### (b). Within-Group Comparisons

Paired-sample *t*-tests were performed to determine participants' progress as recorded by the pre- and post-tests. As can be seen in Table 2, participants' scores improved significantly from the pre-test to the immediate post-test ( $p < .001$ ). Thus, both learning strategies appeared to help them acquire and retain medical terminology in the short term.

While the two groups showed similar results in the first two tests, their scores diverged in the delayed post-test. The collaborative group showed no significant change ( $p = .789$ ) between the immediate post-test ( $M = 36.04$ ,  $SD = 2.67$ ) and delayed post-test ( $M = 36.22$ ,  $SD = 2.67$ ). In other words, participants who engaged in peer interaction continued to display a higher level of vocabulary retention even a month after the treatment, possibly due to deeper cognitive processing.

In contrast, the LLM group's scores dropped significantly ( $p = .012$ ) between the immediate post-test ( $M = 36.62$ ,  $SD = 2.32$ ) and delayed post-test ( $M = 34.20$ ,  $SD = 3.04$ ). The LLM tools thus appeared effective in the short term but failed to achieve the same lasting results as collaborative learning, potentially because these students had less critical discussion.

TABLE 2  
WITHIN-GROUP COMPARISONS

Group	Test Comparison	Mean Scores ( $M \pm SD$ )	<i>P</i>
Collaborative	Pre $\rightarrow$ Post	29.40 $\pm$ 2.38 $\rightarrow$ 36.04 $\pm$ 2.67	< .001*
	Post $\rightarrow$ Delayed	36.04 $\pm$ 2.67 $\rightarrow$ 36.22 $\pm$ 2.67	.789
LLM	Pre $\rightarrow$ Post	29.80 $\pm$ 2.55 $\rightarrow$ 36.62 $\pm$ 2.32	< .001*
	Post $\rightarrow$ Delayed	36.62 $\pm$ 2.32 $\rightarrow$ 34.20 $\pm$ 3.04	.012*

\*Significant at  $p < .05$

###### (c). Between-Group Comparisons

The researcher performed independent-sample *t*-tests in order to compare the groups' overall scores in each of the three tests (see Table 3). This analysis revealed no significant differences between the two groups in the pre-test scores ( $p = .577$ ). In other words, participants' baseline scores for medical terminology, as measured by this test, were roughly equivalent. As with the pre-test, the analysis revealed no significant differences between groups in their post-test scores ( $p = .420$ ). This finding indicated that each of the two strategies employed in this study showed roughly the same level of effectiveness at improving learning outcomes immediately after the treatment.

The trend of rough equivalence found in the pre- and immediate post-tests was broken in the delayed post-test, in which the collaborative group scored significantly higher than the LLM group ( $p = .019$ ). This result could be due to inherent advantages of collaborative learning—such as peer interaction, discussion, and problem-solving—helping students internalize and retain terminology over the long term. Individual LLM-assisted learning, in contrast, might not provide these benefits, or at least not to the same degree.

TABLE 3  
BETWEEN-GROUP COMPARISONS

Comparison	Collaborative Group ( <i>M</i> ± <i>SD</i> )	LLM Group ( <i>M</i> ± <i>SD</i> )	<i>p</i>
Pre-Test	29.40 ± 2.38	29.80 ± 2.55	.577
Post-Test	36.04 ± 2.67	36.62 ± 2.32	.420
Delayed Post-Test	36.22 ± 2.67	34.20 ± 3.04	.019*

\*Significant at *p* < .05

As the results have shown, students showed significant short-term gains in medical terminology acquisition regardless of whether they engaged in collaborative or LLM-assisted translation tasks. Nevertheless, students in the collaborative group showed greater retention of medical terminology a month after the treatment, compared to the significant decline seen with the LLM group.

*B. Qualitative Results*

The six interviewees expressed a variety of opinions about the positive and negative aspects of the different types of tasks. Major themes included differences in student engagement, preferences, and outcomes.

The interviewees from the collaborative group emphasized the benefits of being able to discuss and negotiate with classmates about the translation tasks. According to them, this group work made terminology easier to understand and caused them to be more thoughtful about their translations. One of them said that this type of work was “useful because we explained terms to each other,” and another said it “helped me understand not just the meaning, but why that term is better.” All three of these interviewees stated that collaboration helped their English skills in general and medical vocabulary in particular. For example, one said, “my medical terms improved because we talked through them every time”.

Participants were in favor of implementing more collaborative work in their classes. One of them said, “group work should be in every translation course” because it encouraged a shared sense of responsibility and critical thinking. On the other hand, they noted certain obstacles as well. All of the interviewees claimed they “need more time to discuss and agree” in collaborative tasks and their teacher performed a vital role by guiding discussion and addressing problems. For instance, one of them said that “we needed the teacher to step in when we disagreed or got stuck”.

Another benefit of collaborative learning mentioned by the interviewees was that it boosted their confidence, especially when dealing with new terms. For instance, one said, “I felt more confident when we discussed the answers together.” This suggested that learning in a social environment gave participants more confidence when faced with difficult content. However, they could also get frustrated if someone in their group was less competent, with one participant saying that “sometimes it was slow if someone didn’t understand the task,” an issue that could harm the group’s motivation and progress.

The three interviewees from the LLM group noted how fast and convenient the LLM tools were. For example, one of them said that “it was fast and saved me time,” and another said, “I did the tasks easily, no extra effort needed.” In particular, they remarked about the tools’ effectiveness at dealing with technical vocabulary. One of them, for instance, said, “it’s better than using normal dictionaries.” These interviewees clearly liked the ability to work on their own and said the LLMs resulted in a more manageable and accessible translation process.

On the other hand, these three interviewees also said they sometimes missed having collaborative interaction with peers. One of them said, “sometimes we still need to talk to someone to make sure the meaning is right,” suggesting the need for finding a balance between collaborative and LLM-assisted work. Another concern was relying too much on the LLMs to the point of being a more passive participant in the translation process, as voiced by one of the interviewees: “it made things easy, but I didn’t always think deeply about the terms”.

In addition, although these tools were perceived to work well with simple and intermediate tasks, interviewees were unsure how reliable the tools were in more complicated cases. As one noted, “sometimes the translation sounded good, but I wasn’t sure if it was really accurate.” The lack of certainty expressed in this comment suggested that students were missing out on critical validation when using LLM-generated translations on their own. This would be a particular concern in medical translation, given the need for terminological precision.

V. DISCUSSION

This mixed-methods study examined the relative gains in medical terminology retention among two groups of students: one that worked on translation tasks in collaborative groups and one that worked individually using LLMs. According to the quantitative data, students made significant short-term gains regardless of which group they were assigned to. However, only the students in the collaborative learning group kept those gains a month later, even improving slightly, while the LLM group showed a significant drop in delayed post-test scores.

This was consistent with prior studies that found collaborative learning facilitated higher cognitive processing and engagement (Chyzykova, 2021), supporting language learning over the long term (Ahmadian & Tajabadi, 2020), while LLMs increased efficiency and facilitated vocabulary growth (Kruk & Kaľuzna, 2024). The collaborative group’s gains aligned with Kargar and Ahmadi’s (2021) study, in which structured peer collaboration on translation tasks was related to a greater ability to retain pragmatic and lexical meanings.

The collaborative group's long-term gains agreed with sociocultural theory (Vygotsky, 1978), according to which language is internalized better via social interaction. This finding was also in line with Kiraly's (2000) claim about socio-constructivism that shared meaning-making (rather than passively absorbing information) is beneficial to language learning. In Wang (2019), learners of English as a foreign language showed better retention of vocabulary after engaging in collaborative writing. The collaborative group in the present study was supported by peer negotiation, correction, and discussion, helping students internalize terms.

The LLM group's short-term gains agreed with Yuxiu's (2024) finding that neural machine translation resulted in better translations and higher user satisfaction. This indicated such tools could also be useful in the short term or the early phase of learning.

The significant drop in the LLM group's scores indicated that—despite the utility of ChatGPT and DeepL in accelerating translation—such tools might lead to lower cognitive engagement compared to collaborative learning. Chen (2024) similarly noted that relying too heavily on these tools could lead to a drop in important translation and critical thinking skills. In other words, by seemingly solving translation problems instantaneously, such tools could reduce the opportunities students have for reflecting on problems, testing hypotheses, and contextual reasoning, with a negative impact on terminology retention. In sum, the findings of the present study agreed with Genovese et al. (2024), whose review showed that the usefulness of LLMs was limited in healthcare and similar critical fields; patients in that study liked how fast such tools could make the translation process, but medical personnel were worried about misunderstands due to context. This issue could likewise apply to translation courses, in which translating faster might reduce the quality of education.

The results from the interviews demonstrate a clear distinction between collaborative and individual LLM-Assisted translation activities. These assessments are based on student involvement, learning activities, and confidence in dealing with medical terminology.

The findings revealed that interviewees who participated in collaborative translation projects consistently stressed the value of peer discussions and bargaining in developing a deeper understanding of medical terminology. with interviews indicating that: "My medical term improved because we walked through them every time."; "Group work should be in every translation work." This report is based on current material (Vygotsky, 1978; Storch, 2013). The act of justifying term selections and explaining meanings to others seems to increase metalinguistic awareness, which is an important aspect in vocabulary retention and translation ability. While collaboration promotes engagement and comprehension, it also poses issues in time management and group dynamics.

Interviewees who utilized the LLM-assisted method of learning indicated that: "It is better than using normal dictionaries."; "It made things easy, but I needed to think deeply about what the terms mean." These findings confirm earlier research suggesting that collaborative learning requires planned training to be productive (Dooly, 2008). Participants in the LLM-assisted group valued the speed, convenience, and freedom that language models provided. Their good comments regarding the tools' capacity to handle technical language are consistent with previous research showing that LLMs can help translators by making correct and contextually suitable suggestions (Zhu et al., 2024). Nonetheless, the interviewees also recognized the potential drawbacks of excessive reliance on LLMs. While LLMs can enhance productivity, they may also reduce active engagement and analytical thinking if used without reflection.

## VI. CONCLUSION AND IMPLICATIONS

The results of this study have implications for policy and research. They indicated that collaborative translation tasks could be effective at supporting long-term vocabulary retention, especially when a deeper understanding of the material and terminological precision is necessary, as in medical translation. Despite the potential increase in efficiency provided by LLMs, this study suggests they would better serve the learner by supplementing collaborative learning rather than replacing it. Thus, teachers could apply a blended approach combining these strategies.

Future studies could examine whether these approaches are affected by a student's motivation or language proficiency level, among other variables not analyzed in this study. Researchers could also explore ways to leverage each approach's unique strengths through task design and teacher mediation. As LLMs are used more and more in translation and language education, educators will need to ensure they continue supporting the essential skills learners will need in an evolving global marketplace.

## ACKNOWLEDGEMENTS

I would like to acknowledge all the people who facilitated this project including the research participants for their cooperation in this project.

## FUNDING

This work was funded and supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No: KFU254055).

## REFERENCES

- [1] Ahmadian, M. J., & Tajabadi, H. (2020). Collaborative dialogue, opportunities and challenges in L2 vocabulary learning: An investigation of lexical language-related episodes and vocabulary learning. *Journal of Language and Education*, 6(3), 26–38. <https://doi.org/10.17323/jle.2020.10312>
- [2] Aldossary, K. (2025). Interaction dynamics in collaborative writing: the impact of communication mode and learner goals in EFL dyads. *Theory and Practice in Language Studies*, 15(6), 1834-1845.
- [3] Alshehri, A. A. (2024). Building the Medical Lexicon: A Corpus-Based Approach to Optimising Medical Terminology Acquisition for Pre-Health Science Students. *Forum for Linguistic Studies*, 6(6), pp. 558–574. <https://doi.org/10.30564/fls.v6i6.7400>
- [4] Bayraktar-Özer, Ö., & Hastürkoğlu, G. (2020). Designing collaborative learning environment in translator training: An empirical research. *Research in Language*, 18(2), 137–150. <https://doi.org/10.18778/1731-7533.18.2.02>
- [5] Brahler, C. J., & Walker, D. (2008). Learning scientific and medical terminology with a mnemonic strategy using an illogical association technique. *Advances in Physiology Education*, 32(3), 219–224. <https://doi.org/10.1152/advan.00083.2007>
- [6] Csaba, G., Szabó, I., Környei, J. L., Kerényi, M., Füzési, Z., & Csathó, Á. (2025). Variability in knowledge retention of medical students: repeated and recently learned basic science topics. *BMC Medical Education*, 25(1). <https://doi.org/10.1186/s12909-025-07096-9>
- [7] Dooly, M. (2008). Understanding the many steps for effective collaborative language projects. *Language Learning Journal*, 36(1), 65–78. <https://doi.org/10.1080/09571730801988405>
- [8] Genovese, A., Borna, S., Gomez-Cabello, C. A., Haider, S. A., Prabha, S., Forte, A. J., & Veenstra, B. R. (2024). Artificial intelligence in clinical settings: a systematic review of its role in language translation and interpretation. *Annals of translational medicine*, 12(6), 117.
- [9] Iqbal, L., & Sheeraz, M. (2021). Critiquing Collaborative Language Learning: A Qualitative Study of Teachers' Perceptions. *sjesr*, 4(1), 448–455. [https://doi.org/10.36902/sjesr-vol4-iss1-2021\(448-455\)](https://doi.org/10.36902/sjesr-vol4-iss1-2021(448-455))
- [10] Kargar, A., & Ahmadi, A. (2021). The effect of a collaborative translation task on the learning and retention of pragmatic knowledge. *The Language Learning Journal*, 51(1), 78–93. <https://doi.org/10.1080/09571736.2021.1957989>
- [11] Karwacka, W. (2015). *Medical translation*. ResearchGate. Retrieved January 4, 2025, from [https://www.researchgate.net/publication/314404773\\_Medical\\_translation](https://www.researchgate.net/publication/314404773_Medical_translation)
- [12] Kiraly, D. (2000). *A Social Constructivist Approach to Translator Education: Empowerment from Theory Practice*. Manchester/Northampton: St. Jerome.
- [13] Kornbluth, L., Kaplan, C. P., Diamond, L., & Karliner, L. S. (2021). Communication methods between outpatients with limited-English proficiency and ancillary staff: LASI study results. *Patient Education and Counseling*, 105(1). <https://doi.org/10.1016/j.pec.2021.05.001>
- [14] Kruk, M., & Kałużna, A. (2024). Investigating the role of AI tools in enhancing translation skills, emotional experiences, and motivation in L2 learning. *European Journal of Education*, 60(1). <https://doi.org/10.1111/ejed.12859>
- [15] Kitjaroonchai, N., Kitjaroonchai, T., & Phutiketrakit, C. (2018). Enhancing Thai EFL University Students' English Translation Skills through Online Collaborative Translation. *Abstract Proceedings International Scholars Conference*, 6(1), 295–295. <https://doi.org/10.35974/isc.v6i1.1291>
- [16] Montalt-Resurrecció, V., & Shuttleworth, M. (2021). Research in translation and knowledge mediation in medical and healthcare settings. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 11. <https://doi.org/10.52034/lansstts.v11i.294>
- [17] Mossop, B., Hong, J., & Teixeira, C. (2019). *Revising and Editing for Translators*. Routledge. <https://doi.org/10.4324/9781315158990>
- [18] Noll, R., Berger, A., Kieu, D., Mueller, T., Bohmann, F. O., Müller, A., Holtz, S., Stoffers, P., Hoehl, S., Oya Guengoeze, Eckardt, J.-N., Holger Storf, & Schaaf, J. (2025). Assessing GPT and DeepL for terminology translation in the medical domain: A comparative study on the human phenotype ontology. *BMC Medical Informatics and Decision Making*, 25(1). <https://doi.org/10.1186/s12911-025-03075-8>
- [19] Olvera-Lobo, M., Robinson, B., Senso, J. A., Muñoz Martín, R., Muñoz-Raya, E., Murillo-Melero, M., Quero-Gervilla, E., Castro-Prieto, M., & Conde-Ruano, T. (2009). Teleworking and collaborative work environments in translation training. *Babel*, 55(2), 165–180. <https://doi.org/10.1075/BABEL.55.2.05OLV>
- [20] Özyurt, S. (2024). AI-assisted English language learning for cross-cultural medical education in multilingual settings. *Experimental and Applied Medical Science*, 5(2). <https://doi.org/10.46871/eams.1464830>
- [21] Patil, S., & Davies, P. (2014). Use of Google Translate in medical communication: Evaluation of accuracy. *BMJ*, 349(1), g7392–g7392. <https://doi.org/10.1136/bmj.g7392>
- [22] Rao, P., McGee, L. M., & Seideman, C. A. (2024). A Comparative assessment of ChatGPT vs. Google Translate for the translation of patient instructions. *Journal of Medical Artificial Intelligence*, 7, 11. <https://doi.org/10.21037/jmai-24-24>
- [23] Ratna, H. (2019). *The Importance of Effective Communication in Healthcare Practice - BCPHR Journal*. BCPHR. Retrieved March 18, 2025 from <https://bcphr.org/23-article-ratna/>
- [24] Saeed, M., Ahmed, L., AbdAlla, E., Zinab Alatawi, Alhowiti, A. M., Tasneem S. A. Elmahdi, Mohammed, S., & Elhag, A. (2025). Retention of gross and clinical anatomy knowledge among medical graduates in Sudan: a comparative study. *BMC Medical Education*, 25(1). <https://doi.org/10.1186/s12909-025-06832-5>
- [25] Sánchez Ramos, M. D. M. (2019). Mapping new translation practices into translation training. *Babel*, 65(5), 615–632. <https://doi.org/10.1075/babel.00114.san>
- [26] Storch, N. (2013). Collaborative Writing in L2 Classrooms. *New Perspectives on Language & Education*, 31. <https://doi.org/10.21832/9781847699954>
- [27] Vygotsky, L. S. (1978). *Mind in Society the Development of Higher Psychological Processes*. Cambridge, MA Harvard University Press.

- [28] Wang, W. (2019). EFL incidental vocabulary acquisition and retention through collaborative written output tasks. *RELC Journal*, 50(1), 61–75. <https://doi.org/10.1177/0033688217730140>
- [29] Xu, J. (2024). *New Thoughts on Translation*. *Qizhen Humanities and Social Sciences Library*. Singapore: Springer Nature Singapore. <https://doi.org/10.1007/978-981-97-6386-3>
- [30] Yuxiu, Y. (2024). Application of Translation Technology based on AI in Translation Teaching. *Systems and Soft Computing*, 6, 200072–200072. <https://doi.org/10.1016/j.sasc.2024.200072>
- [31] Zhu, D., Chen, P., Zhang, M., Haddow, B., Shen, X., & Klakow, D. (2024). Fine-Tuning Large Language Models to Translate: Will a Touch of Noisy Data in Misaligned Languages Suffice? In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (pp. 388-409). Association for Computational Linguistics.



**Khaled Aldossary** is currently an Associate Professor of Applied Linguistics and the Head of English Department, King Faisal University, Al Ahsa, Saudi Arabia. He received his Ph.D. in Applied Linguistics and Educational Linguistics from the University of Melbourne in September 2017. His academic interest areas are ESL/EFL academic writing, feedback on writing, collaborative learning, and computer-mediated language learning. Affiliation: Department of English Language, King Faisal University, Al Ahsa, Saudi Arabia. E-mail: [kaldossary@kfu.edu.sa](mailto:kaldossary@kfu.edu.sa) Phone: +966566595868 ORCID iD: <https://orcid.org/0000-0002-4917-3258>.