ISSN 1799-2591 (print); ISSN 2053-0692 (online)
Theory and Practice in Language Studies, Vol. 16, No. 3, pp. 732-740, March 2026
DOI: https://doi.org/10.17507/tpls.1603.03

AI Versus Human Lexicographers: A Comparative Analysis of Translation Strategies for Arabic Collocations and Cultural References*

Said Faiq
American University of Sharjah, United Arab Emirates

Abstract—**Heliel (2000) identified a set of Arabic collocations and cultural references that have consistently challenged human lexicographers, with many opting to simply omit them from their compilations, including the widely used *Al-Mawrid Arabic-English dictionary*. Deploying Pedersen's (2011) taxonomy of translation strategies, this study interpretatively evaluated 150 English translations of this set. The translations were mined from three Arabic-English bilingual dictionaries compiled by human lexicographers (Baalbaki, 2001; Abu-Ssaydeh, 2013; Hafiz, 2004) and two sets were generated by the two leading artificial intelligence (AI)-powered systems: ChatGPT and Google Gemini. The findings reveal a striking contrast in strategy use: while human lexicographers frequently omitted difficult phrases, AI tools provided complete translations for all expressions (100%). Despite their relatively recent development and launch, the two AI-systems exhibited lexicographical capabilities comparable to human lexicographers, particularly in adopting target-oriented strategies. This suggests that such tools could complement traditional lexicography by enhancing coverage, efficiency, and contextual richness. Perhaps, a hybrid "H-AI" approach may well be the way forward.**

Index Terms—**artificial intelligence, human lexicography, Arabic to English translation, ChatGPT, Google Gemini**

I. Introduction

As communication, translation brings cultures together through often-manipulative negotiations of cultural differences whereby sources are usually adapted to reflect prevailing cultural norms of the target cultures (Baker, 2018). This becomes particularly problematic when translating collocations and cultural references, which frequently lack direct equivalents in the target language—especially between linguistically and culturally distant languages such as Arabic and English. In such cases, bilingual dictionaries serve as essential resources, underscoring the interdependence of lexicography and translation.

Given its importance for translation, lexicography significantly influences quality. Bilingual dictionaries are often the translator's (and other language users) first port of call, but their limitations in addressing collocations and cultural references tend to hinder rather than help the translation process. Translators frequently need to move beyond the literal meanings provided in bilingual dictionaries to preserve the original context and ensure culturally appropriate renderings. Unfortunately, most bilingual dictionaries offer inadequate support for culturally sensitive translations (Abu-Ssaydeh, 2012). This issue is especially evident in Arabic to English translation, where collocations and cultural references pose significant challenges, particularly those deeply couched in religious, social, or historical contexts. Here, achieving effective translations requires not only linguistic gymnastics but also cultural awareness and approximation to ensure that translations are fully relevant to both source and target languages and cultures.

Simply put, collocations are word pairings that naturally co-occur in a language and often carry fixed meanings. Translating these specific language expressions presents unique difficulties due to their semantically restricted combinations and specific connotations. For example, 'deliver a speech,' 'deliver a pizza' or 'deliver a baby' illustrate how the same verb collocates differently depending on intended yet unique and fixed meanings (Baker, 2011). Similarly, cultural references encapsulate a community's history, religion, values, and social structures and practices. These language items are challenging as they tend to cause misinterpretations and, consequently, mistranslations.

Traditional Arabic-English bilingual dictionaries often present words in isolation without contextual cues required to capture embedded meanings. To address this gap, supplementary resources—such as glossaries of cultural terms, encyclopaedias, and even consultations with native speakers or subject-matter experts—are necessary, particularly when dealing with complex collocations and cultural references. Abu-Ssaydeh (2012), for example, critiques the limitations of Arabic-English dictionaries for not effectively representing collocations and suggests systematic improvements, such as refining collocational ranges and enhancing metaphorical representations to better support learners and translators of Arabic.

In an increasingly globalized world, the demand for translation has grown exponentially. To meet this demand, machine

* This paper represents the opinions of the author(s) and does not mean to represent the position or opinions of the American University of Sharjah.

translation (MT)—using computational algorithms to translate text without human intervention—has been widely adopted, driven mostly by economic imperatives (Hassan et al., 2018; Alkhatib & Shaalan, 2018). Since its early adoption, MT has undergone significant evolution, transitioning from early rule-based approaches (relied on grammar rules and dictionaries) through statistical methods, and more recently, to Neural Machine Translation (NMT). The latter has revolutionized the field by leveraging neural networks to produce more fluent translations (Koehn & Knowles, 2017) and paved the way for the recent introduction of artificial intelligence (AI) systems for translation. The shift from traditional human translation through MT to AI-powered tools has reshaped the translation academic and business landscapes.

Research on Arabic-English lexicography points to the conclusion that despite advances in machine and AI translation, human translation remains more reliable for effectively rendering Arabic collocations and cultural references, whereby they explore contextual contours necessary to interpret and render such expressions often innovatively, while both preserving original meanings and ensuring relevance in the target language. While MT systems are easily accessible and mostly cost-effective, they often struggle with culturally embedded expressions as they "lack sensitivity to idiomatic language and cultural context, resulting in translations that are structurally sound but semantically flawed" (Koehn & Knowles, 2017, p. 72). Although recent advancements in AI-powered translation have improved the handling of idiomatic language uses, they generally still struggle with highly contextual sensitive human language use such as collocations and cultural references, particularly for a language such as Arabic whereby linguistic, cultural, and contextual subtleties are critical for translation (Eddahibi & Mensouri, 2023).

The early launched AI systems like ChatGPT (GPT), released in November 2022, and Google Gemini (GG), launched in December 2023, have made remarkable progress in handling context and adapting translations through reinforcement learning from human feedback (RLHF) (Alkhawaja, 2024). However, their ability to interpret deeply embedded cultural references remains limited (Jiao et al., 2023). By incorporating deep learning and generative capabilities, AI systems have demonstrated notable improvements in contextual and stylistic interpretation as they continue to refine their ability to differentiate between literal and idiomatic meanings, thereby enhancing translation appropriateness. Still, researchers have reported that these tools struggle with highly specific cultural and idiomatic expressions and produce lower quality translations than humans. Hao et al. (2025, p. 45) aptly comment:

> Despite their impressive capabilities in interacting with humans and handling various natural language processing tasks, dialogue-based large language models like ChatGPT may still provide unsatisfactory responses in certain conversational scenarios. This is because these models are based on generative models that rely on statistical patterns in the data they were trained on, rather than in-depth understanding or true comprehension of the content.

As such, human translators remain largely the benchmark for quality since they "draw upon cultural and contextual awareness to convey the essence of the message," as Munday (2016, p. 43) points out. This is a skill AI systems have yet to fully replicate but are fast developing through advanced algorithms and RLHF. In Arabic-English translation, collocations and cultural references require special handling that only human translators are said to provide, as they engage in interpretative processes that ensure effective renderings. This highlights the inseparability of language and culture in translation—whereby culture shapes language, and language, in turn, reflects culture and where even humans often face challenges in effectively navigating this interwoven relationship (Faiq, 2019).

Building on this premise and on Heliel's (2000) findings regarding the challenges posed by the 30 problematic Arabic expressions (collocations and cultural references), this study investigates two hypotheses:

1. Given their reported limitations, AI-powered translation tools of GPT and GG will omit collocations and cultural references when translating from Arabic into English.
2. AI-powered translation tools of GPT and GG will face difficulties in interpreting and effectively rendering collocations and cultural references from Arabic into English.

To evaluate these two hypotheses, this study compared 150 translations of the Arabic collocations and cultural references delimited by Heliel (2000). In so doing, the study evaluates the performance of two widely used AI translation tools—GPT and GG—against three leading Arabic-English bilingual dictionaries authored by Arab lexicographers (human lexicographers (HLs)). The comparison is based on Pedersen's (2011) typology of translation strategies supplemented by two strategies used for this study. The remainder of this study is structured as follows: section 2 outlines the material and method; section 3 presents the results followed by the discussion of these results in section 4. The conclusion provides a summary and some suggestions for further research.

II. Material and Method

This study evaluates 150 English translations of the 30 challenging Arabic collocations and cultural references identified by Heliel (2000) as most challenging even for human lexicographers, let alone AI-systems. As shown in Appendix, three sets of translations were extracted from Arabic-English bilingual dictionaries, while two sets were generated by ChatGPT (GPT), GPT-4 model, and Google Gemini (GG), Gemini 2.5 Pro model; both on June 13, 2024. The prompt used for each system was: 'taking context into consideration, provide a translation of the uploaded Arabic document into idiomatic English'.

Launched in November 2022, GPT is designed by OpenAI and uses transformer-based models and RLHF to enhance conversational capabilities. GPT is continuously improving natural language understanding and generation. Released over

a year later (December 2023), GG is an AI model developed by Google DeepMind. It combines Google's advanced NLP capabilities with RLHF and deep neural network integration. It is designed to offer a competitive AI alternative, particularly to GPT. Both GPT and GG use transformer architectures and reinforcement learning from human feedback. Both are designed to handle complex language processing tasks to generate natural, fluent, and context-aware texts for a wide range of applications, including translation. Choosing between these two large language models (LLMs) depends on the purpose of use (GPT offers creativity and conversational depth, while GG offers precision and factual accuracy). In an ideal world, the combination of the strengths of both LLMs through a hybrid approach could lead to more robust and versatile AI applications.

The sets of data by human lexicographers (HLs) were sourced from three leading Arabic-English bilingual dictionaries:

1. Human Lexicographer 1 (HL1): *Al-Murshid Arabic-to-English Dictionary* (MUR) is compiled by Abu-Ssaydeh (2013). This bilingual dictionary provides translations and usage notes with a particular emphasis on assisting users in handling culturally embedded collocations and idiomatic expressions.
2. Human Lexicographer 2 (HL2): *Al-Mawrid Arabic-English Dictionary* (MAW) is compiled by Baalbaki (2001). Widely regarded for its extensive vocabulary and user-friendly format, this dictionary is a primary reference for Arabic-English translation (the 737 of Arabic-English dictionaries, so to speak). However, as noted by Heliel (2000), its 1988 edition offered limited treatment of culturally specific references and collocations and included only 17% of the 30 phrases. Its 2001 edition (fourteenth) is not much better either as discussed in this study.
3. Human Lexicographer 3 (HL3): *Al-Hafiz Arabic Collocations Dictionary* (HAF) is compiled by Hafiz (2004). Unlike the previous dictionaries, this resource is dedicated exclusively to Arabic collocations and their English equivalents.

The translations provided by these three HLs along with the two sets generated by the AI tools (GPT and GG) were analyzed both quantitatively and qualitatively using Pedersen's (2011) taxonomy of translation strategies to which two strategies were added for the purpose of this study. The full datasets, including the identified strategies for each of the 150 translations, are presented in Appendix. The identification / categorization of strategies was conducted by the author of this study, drawing on nearly four decades of academic experience and complemented by Fayed's (2007) monolingual (Arabic-Arabic) dictionary of collocational expressions.

Pedersen (2011) offers a comprehensive framework for assessing the translation of cultural references and collocations. The taxonomy includes six primary strategies: source-oriented strategies (SO) or target-oriented strategies (TO):

- Source-Oriented Strategies (SO)
- Retention (R): Here, the cultural reference is retained in its original form or slightly adapted to fit the target language. An R translation may appear italicized or otherwise marked to indicate its foreign origin.
- Specification (SP): This strategy is usually materialized through the insertion of additional information or the inclusion of further contextual clues to clarify the cultural reference and enhance comprehension.
- Direct Translation (DT): Through this strategy, cultural reference is translated directly, without modification or added explanation. Except for some necessary language adjustments, the structure and meaning are preserved as much as possible as with proper and brand names, for example.
- Target-Oriented Strategies (TO)
- Generalization (G): Here, the source element is replaced with a broader / general term in the target language that maintains the overall meaning but loses the specificity of the source. This strategy often yields less precise translations.
- Substitution (SU): This strategy results in the replacement of the source item with a corresponding reference in the target culture; often involving a completely different rendering.
- Omission (O): Here the source items are simply omitted (deleted, excluded) if they are deemed unnecessary or too complex to translate.
- Other Strategies (OS)

While Pedersen's framework provides a structured approach, real-world translation often requires other strategies to address complex translational tasks. To account for this, an additional category—other strategies (OS)—is suggested in this study, encompassing two strategies:

- Mistranslation (MT) refers to instances where translations were simply incorrect.
- Combination of Strategies (CS) refers to cases where two or more translation strategies were applied to render a single collocation or cultural reference.

By applying these strategies, this study assesses how GPT and GG compare with the three HLs in handling the translation of the challenging Arabic collocations and cultural references into English. The following section presents the results of this analysis.

III. Results

Appendix provides the full datasets with the strategies assigned to each of the 150 translations by the two AI systems (GPT and GG) and the three human lexicographers (HL1, HL2, HL3). Table 1 provides the total translations and percentages for each translator (t=total).

TABLE 1
TOTAL TRANSLATIONS FOR GG, GPT AND HLS

Arabic source phrases	GG		GPT		HL1 (MUR)		HL2 (MAW)		HL3 (HAF)	
30	t	%	t	%	t	%	t	%	t	%
	30	100	30	100	20	67	5	17	3	10

Table 2 provides a summary of the translation strategies assessment and frequencies per translator (t=total).

TABLE 2
SUMMARY OF TRANSLATION STRATEGIES ASSESSMENT FOR GG, GPT AND HLS

Strategy		GG		GPT		HL1 (MUR)		HL2 (MAW)		HL3 (HAF)	
		t	%	t	%	t	%	t	%	t	%
SO	R	0	0	0	0	0	0	0	0	0	0
	SP	0	0	1	3.3	4	13.3	0	0	0	0
	DT	18	60	15	50	8	26.7	0	0	0	0
TO	G	1	3.3	2	6.7	4	13.3	1	3.3	1	3.3
	SU	3	10	1	3.3	3	10	1	3.3	2	6.7
	O	0	0	0	0	10	33.3	25	83.3	27	90
OS	MT	5	16.7	0	0	0	0	1	3.3	0	0
	CS	3	10	11	36.7	1	3.3	2	6.7	0	0
Total		30		30		30		30		30	

Table 3 outlines the total frequencies and reflective percentages for all strategies for the five translators (GG, GPT, and the 3 HLs).

TABLE 3
TRANSLATION STRATEGIES USED AND PERCENTAGES FOR GG, GPT AND HLS

Strategy		Total	%
SO	R	0	0%
	SP	5	3.3%
	DT	41	27.3%
Total		**46**	**30.61%**
TO	G	9	6%
	SU	10	6.7%
	O	62	41.3%
Total		**81**	**54%**
OS	MT	6	4%
	CS	17	11.3%
Total		**23**	**15.3%**
Total (all)		150	100%

The results provide valuable insights into the strategies employed by the two AI-powered tools (GPT and GG) and the three human lexicographers (HL1, HL2, and HL3) in producing 150 translations into English of the 30 complex Arabic collocations and cultural references. Pedersen's (2011) taxonomy of translation strategies along with the two supplementary strategies—mistranslation (MT) and combination of strategies (CS) specifically used in this study; were deployed in this study. As Tables 2 and 3 show, source oriented (SO) strategies, which aim to retain or minimally alter the original expressions, were scarcely used across all translators. Specifically, retention (R) was entirely absent (0%) in all 150 translations. This finding may underscore the inherent challenges of directly transferring Arabic collocations and cultural references into English without some adaptation and even manipulation. This finding also highlights the interesting point that the two AI-tools behaved like the three HLs by not attempting to retain/preserve collocations and cultural references in their original form. Tables 2 and 3 also indicate that target oriented (TO) strategies, which prioritize the expectations of the target audience (domestication), were the most employed, accounting for 54% of all uses. The two other strategies of MT and CS accounted for 15.3% of all uses, reflecting both linguistic and cultural challenges as well as innovative ways in rendering Arabic collocations and cultural references into English, particularly the use of CS tools.

IV. Discussion

The interpretative assessment of the 150 translations produced by three human lexicographers (HLs) and two AI systems (GPT and GG) reveals that TO strategies dominated the dataset, being employed nearly twice as SO strategies. Table 3 shows that SO strategies were used 46 times out of 150 (30.6%), with GG slightly surpassing GPT at 18/30 uses (60%) compared to 16/30 (53.3%). Among the three HLs, only HL1 employed SO strategies 12 times (40%), while HL2 and HL3 did not use them at all.

Conversely, TO strategies accounted for 81 out of 150 uses (54%). The three HLs exhibited the highest use of TO strategies, with HL1, HL2, and HL3 employing them 17, 27, and 30 times, respectively. Meanwhile, GPT and GG used TO strategies sparingly, with only 4 and 3 occurrences each. Notably, omission (O) was the most frequently used TO strategy, appearing 62 times (41.3%), almost exclusively by HLs, with HL3 using it most (27 times, 90%), followed by HL2 (25 times, 83.3%) and HL1 (10 times, 33.3%). The prevalence of omission among HLs points to their tendency to systematically exclude any of the Arabic phrases that they deemed to be complex or those that indicate highly culturally specific references and content. In contrast, GPT and GG did not use omission at all, reflecting their algorithmic design to produce complete translations. This finding contradicts hypothesis 1 of this study, which posited that unlike AI systems, human lexicographers (translators) would at least provide explanatory renderings for Arabic collocations and cultural references in English to assist translators. The discussion below examines Pedersen's two categories (SO and TO), alongside the category of OS, which was introduced specifically for this study.

A. Source-Oriented Strategies (SO)

The limited use of SO strategies highlights the challenge of preserving the meanings of Arabic collocations and cultural references. Notably, retention (R) was not used at all, suggesting that all translators—both human and AI—prioritized finding appropriate target language renderings that approximated the source meanings in English. This complete absence of R reflects the difficulty of maintaining the semantic and cultural weight and complexity of the Arabic phrases analyzed here as underlined by Heliel (2000).

Specification (SP) appeared only five times (3.3%) across the 150 translations, as shown in Table 3, with GPT employing it once and HL1 four times. The infrequent use of SP reflects the difficulty of adding explanatory information to translations, a task that often requires subtle contextual understanding—something AI tools may still be developing. Examples 3 (أم القرى) and 29 (بيوت الله), Appendix, indicate how SP was used by GG (Mecca), GPT (Mecca) and HL1 (Mekka) in their English translations. HL2 and HL3 simply omitted these two examples.

Direct Translation (DT), the most frequently used SO strategy, appeared 41 times (27.3%), as shown in Table 3. GPT and GG relied heavily on DT, using it 18 and 15 times, respectively, compared to eight uses by HL1. HL2 and HL3 did not use DT at all. The preference for DT by AI tools suggests their tendency to adhere closely to the semantic loads of the source Arabic phrases. HLs often deviated from DT, instead employing TO strategies like substitution (SU) to achieve greater cultural relevance. For instance, in rendering item 10 (خفف الوطأ, Appendix), GG, GPT and HL1 used DT in their translations to produce 'lightened the burden,' 'eased the pressure,' and 'to lighten the burden,' respectively. Once again, HL2 and HL3 simply opted to omit.

B. Target-Oriented Strategies (TO)

As Table 3 shows, the reliance of the three HLs on TO strategies demonstrates, in theory, their ability to navigate the complexities of the target language and culture. Among TO strategies, substitution (SU) was the most used after omission (O), occurring 10 times (6.7%). The three HLs accounted for six of these SU uses, while GPT and GG used it three times and once, respectively. The use of SU often requires cultural adaptation, showcasing some cultural gymnastics necessary to translate specific collocations and cultural references effectively. For example, item 26 (بيت دعارة, lit. house of adultery, fornication, immorality, prostitution), Appendix, was rendered as 'brothel,' 'a brothel,' brothel,' and 'brothel' by GG, GPT, HL2 and HL3, respectively. Strangely enough, HL1 excluded (omitted) this expression.

Generalization (G), which simplifies cultural references for accessibility at the cost of some nuance, was used nine times (6%). HL1 applied it most frequently (four times), while the other two HLs used it only once each. GPT and GG applied G sparingly, with one and two uses, respectively. For example, GPT, HL1, HL2 and HL3 all used G to render item 13 (رفع الكلفة), Appendix, as removed formalities, to take liberties, take liberties with, and to take liberties with, respectively. GG mistranslated (MT) this item because it misinterpreted the intended connotations focusing on the literal meaning instead.

Omission (O) emerged as the most dominant strategy overall, accounting for 41.3% of all strategies. Human translators (three HLs), particularly HL2 (83.3%) and HL3 (90%), relied heavily on omission, frequently excluding many Arabic collocations and cultural references from the dataset they most likely considered too obscure or challenging. On the other hand, GPT and GG did not use omission at all, reflecting their algorithmic design to produce full translations regardless of difficulty. This contrast underscores differing priorities: while human translators may prioritize readability and relevance, AI tools strive for comprehensiveness.

The high frequency of omission by HL2 and HL3 and the complete absence of omission among the AI tools (GPT and GG) challenge conventional assumptions in translation studies. At the very least, GPT and GG provided renderings that could assist translators and other users in deciphering the cultural and idiomatic meanings of the Arabic expressions. HL1 performed better than HL2 and HL3, but still employed omission 10 times, which is not ideal for translators and others seeking guidance with challenging linguistic structures. Regardless of quality, by generating renderings for the entire dataset (100%), both GPT and GG offer insights and starting points for translators handling such difficult expressions. This observation undermines and challenges the second hypothesis set out for this study.

C. Other Strategies (OS)

As Table 3 shows, the supplementary strategies of mistranslation (MT) and combination of strategies (CS) accounted for 15.3% of all strategies. MT occurred six times (4%), predominantly in GG's output (five times, 16.7% for items 4, 8, 9, and 13), with one instance by HL2 (3.3%, for item 16). GPT, HL1 and HL3 produced no MTs. Even in the case of the five mistranslations (MTs) produced by GG (and one by HL2), translators and other users could still derive some help (at least identify inappropriate renderings and potentially reverse engineer them into effective translations). That HL2, a human lexicographer, produced an MT was unexpected.

Combination of strategies (CS) was used 17 times (11.3%), with GPT accounting for the majority (11 cases, 36.7%). GG used CS three times (10%), while HL1 and HL2 applied it once (3.33%) and twice (6.7%), respectively. For example, item 2 (أهل اللغة, Appendix) was respectively rendered by GPT and GG as linguists, people of language and the people of language, linguists. Both GPT and GG used DT and SU, albeit not in the same order, to handle this item through translation into English.

The minimal use of CS by the three HLs—once and twice by HL1 and HL2, respectively—compared with the 17 occurrences by GPT and three by GG is significant. This suggests that prevailing assumptions about the inability of AI-powered translation tools to handle context-sensitive Arabic expressions, in this case, may now be outdated. By using CS to generate English translations, GPT appears to recognize the value of offering multiple alternatives or additional contextual information; an approach HLs did not adopt in their translations.

The analysis reveals distinct patterns in strategy use between GPT, GG and the three HLSs:

- GPT and GG excelled in producing complete translations, avoiding omissions while heavily relying on DT. However, GG struggled with some cultural meanings and ended up generating 5 MTs. GPT's higher use of CS and its zero MT renderings suggest an emerging (and being fast augmented) ability to handle complex translation tasks such as collocations and cultural references even from difficult languages, including Arabic.
- HL1, HL2 and HL3 demonstrated the expected sensitivity to cultural and contextual subtleties. This is axiomatically reflected in their use of mostly target oriented strategies. However, their heavy reliance on omission (O), especially HL2 and HL3, suggests a tendency to simplify translations by excluding difficult phrases, which may compromise textual completeness and is not helpful for translators, learners and other users of Arabic.

The findings of this study largely align with existing literature on translating collocations and cultural references while simultaneously challenging certain assumptions about some limitations of AI (Eddahibi & Mensouri, 2023; Alowedi et al., 2023). El-Farahaty and Alwazna (2024), for example, point out the low use of target oriented (TO) strategies in subtitling; a finding that contrasts with the heavy reliance on TO strategies by the three HLs reported in this study. Future AI developments should focus on improving cultural adaptability, and human lexicographers might benefit from leveraging AI tools to balance comprehensiveness, efficiency, and cultural appropriateness.

V. Conclusion

Based on the received wisdom and documented research on AI-powered translation tools, this study set out to test two hypotheses: GPT and GG would omit (delete) Arabic collocations and cultural references when translating into English and would most likely face difficulties in interpreting and effectively translating such Arabic expressions. The results of this study challenge both hypotheses. They reveal that the two AI tools of GPT and GG managed to generate translations that could easily compete with human lexicographers (HL1, HL2 and HL3). Contrary to expectations, AI-tools did not employ omission (deletion) at all as a strategy, whereas the three HLs omitted a considerable number of collocations and

cultural references, up to 83% and 90% of the dataset was omitted by HL2 and HL3 respectively. GPT and GG produced complete translations, offering interpretable renderings that, while not precise in rare cases generated by GG, still provide valuable insights for human translators and other language users. At the same time, hypothesis 2 remains partially valid. The results confirm that AI-translation tools still face difficulties in capturing cultural and idiomatic meanings, often defaulting to literal DT or G instead of employing SU or SP, which require deeper cultural adaptation. But even here, both did not generate the usual howlers associated with MT systems like google translate. Furthermore, the three HLs omitted far more of the Arabic expressions when GPT and GG omitted zero. The overuse of O by the three HLs is not helpful for all those who need such information, including 'poor' translation students or general language learners.

This study underscores the evident complementary nature of AI and human lexicographers for translation and other domains and that a hybrid model combining AI with human lexicography could optimize translation quality of bilingual dictionaries. The interplay between AI and human lexicography underscores their complementary strengths and paves the way for more effective human-AI collaboration in translating culturally specific content. While AI excels in completeness and consistency, human lexicographers remain better at contextual and cultural sensitivity. GPT demonstrated greater flexibility than GG, employing CS more frequently to offer culturally relevant renderings and explanations. For example, research should explore interactive AI-assisted lexicography from Arabic to English that would allow human translators and users to edit, adapt, and supplement AI-generated output, ensuring both appropriateness and cultural resonance while leveraging the computing powers of AI-systems.

Ultimately, translation is more than a linguistic task, it is a cultural exchange. As AI continues to evolve, its role should, hopefully, be to support, rather than replace, human lexicographers and translators, ensuring that cross-cultural communication remains effectively and appropriately meaningful in an interconnected world where AI is set to play an increasingly prominent role. A role humans need to accept and learn to adapt to and adopt in all walks of life, including translation and associated disciplines such as lexicography. The integration of context-sensitive neural networks and culturally adaptive translation strategies could significantly bridge the gap between AI-generated and human-molded translations, leading, for example, to effective, appropriate, and useful Arabic to English lexicographical resources.

APPENDIX. ENGLISH TRANSLATIONS (GG, GPT, HL1, HL2, HL3) AND STRATEGIES EMPLOYED

1) **R** = Retention
2) **SP** = Specification
3) **DT** = Direct translation
4) **G** = Generalization
5) **SU** = Substitution
6) **O** = Omission
7) **MT** = Mistranslation
   ST = Strategy
8) **CS** = Combination of strategies

No.	Arabic Source Text	GG		GPT		HL1 (MUR)		HL2 (MAW)		HL3 (HAF)	
		Translation	ST	Translation	ST	Translation	ST	Translation	ST	Translation	ST
1	رأس الحكمة	The pinnacle of wisdom	G	The pinnacle of wisdom	G	Fear of God is the beginning of wisdom	SP		O		O
2	أهل اللغة	Linguists, people of language	SU/DT	The people of language / Linguists	DT/SU		O		O		O
3	أم القرى	Mother of cities (often referring to Mecca)	DT/SP	The mother of all cities (Mecca)	DT/SP	Makka	SP		O		O
4	مناط الأمل	The foundation of hope	MT	The source of hope	DT		O		O		O
5	إخوة السلاح	Brothers in arms	DT	Brothers in arms	DT	Comradeship (-in-arms)	SU		O		O
6	فتيل الحرب	The spark of war	DT	The spark of war	DT		O		O		O
7	علت أسهمه	His stock has risen	DT	His stocks have risen / His value increased	DT/SP		O		O		O
8	مزق شملهم	They were divided	MT	Tore them apart	DT	Rend	G		O		O
9	فرق شملهم	They were separated	MT	Dispersed their unity	DT	Drive a wedge between	SU		O		O
10	خفف الوطأ	Lightened the burden	DT	Eased the pressure	DT	To lighten the burden	DT		O		O
11	خفف من غلوائه	Calmed his anger	DT	Moderated his intensity / Calmed him down	DT/SP		O		O		O
12	وحد الكلمة	United in word	MT	United the stance / Unified the word	SP	To unify; to unite	DT		O		O
13	رفع الكلفة	Increased the cost	MT	Removed formalities	DT	To take liberties	G	To make free with Take liberties with	G	To be free with; to take liberties with	G
14	دق طبول الحرب	Beat the drums of war	DT	Beat the drums of war	DT	Beating (the drums)	DT		O		O
15	رجحت الكفة	The scale tipped	DT	The balance tipped	DT	To tip the balance / scales (for)	DT		O		O
16	قميص عثمان	A contentious issue, a bone of contention	SU	A pretext for conflict (based on a historical event)	SP	A cause for controversy	SU	Bloody shirt	MT		O
17	شعرة معاوية	A very fine distinction, a hair's breadth	SU	The fine thread of diplomacy / Strategic patience	G		O		O		O
18	وأد الفكرة في مهدها	Stifled an idea in its infancy	DT	Killing the idea in its cradle	DT		O		O		O
19	خطاب مفتوح	An open letter	DT	An open letter	DT	Open letter	DT		O		O
20	سلقه بلسانه	Scorched him with his tongue	DT	Lash him with words / Verbally attacked him	DT/SP	To scald / smite with one's tongue	DT		O		O
21	ابن المدينة	A native of the city	DT	A city native	DT	Native of a place	G		O		O
22	أبناء الوطن	Sons of the nation	DT	Sons of the nation / Citizens	DT/G		O		O		O
23	ابن الخطيئة	Son of sin	DT	A child of sin (usually refers to an illegitimate child)	DT/SP	Bastard	G		O		O
24	ابن السبيل	A traveler, a wayfarer	DT	A wayfarer / Traveler in need	DT/SP	A traveler or Wayfarer	DT	Wayfarer, traveler, passer-by, walker, vagabond, wanderer, tramp, hobo, vagrant	DT/SP		O
25	بيت الشباب	Youth hostel	DT	A youth hostel	DT	Youth hostel	DT		O		O
26	بيت دعارة	Brothel	SU	A brothel	SU		O	brothel, bordello, bawdy house, whorehouse, house of prostitution	SU	Brothel	SU
27	بيت المسنين	Nursing home	DT	A retirement home / Elderly care home	DT/SP		O		O		O
28	بيت الزوجية	Marital home	DT	The marital home	DT	Marital house; conjugal home	DT/SU	Marital house, conjugal home	DT/SU	Conjugal home	SU
29	بيوت الله	Houses of God (mosques)	DT/SP	Houses of God (usually refers to mosques)	DT/SP	Mosques	SP		O		O
30	بدد الأوهام	Dispelled illusions	DT	Dispelled illusions	DT	To drive away (bad thoughts)	SP		O		O

REFERENCES

[1] Abu-Ssaydeh, A. F. (2013). *Al-Murshid: A General Arabic-English dictionary*. Dar Asharq wa Algharb.

[2] Abu-Ssaydeh, A. F. (2012). Lexicalized Metaphors in the Bilingual Dictionary: The Case of A-E-A Dictionaries. *Turjuman*, *21*(1), 11–53.

[3] Alkhatib, M., & Shaalan, K. (2018). Paraphrasing Arabic Metaphor with Neural Machine Translation. *Procedia Computer Science*, *142*(ACLing), 308–314.

[4] Alkhawaja, L. (2024). Unveiling the New Frontier: ChatGPT-3 Powered Translation for Arabic-English Language Pairs. *Theory and Practice in Language Studies*, *14*(2), 347-357. https://doi.org/10.17507/tpls.1402.05

[5] Alowedi, N., Mohammed, A. A., & Al-Ahdal, H. (2023). Artificial Intelligence based Arabic-to-English machine versus human translation of poetry: An analytical study of outcomes. *Journal of Namibian Studies*, *33 S2*, 1523–1538. DOI: 10.59670/jns.v33i.800

[6] Baalbaki, R. (2001). *Al-Mawrid. A Modern Arabic-English dictionary* (14th ed.). Dar El-Ilm Lilmalayin.

[7] Baker, M. (2018). *Translation and conflict: A Narrative Account*. Routledge.

[8] Baker, M. (2011). *In other words: A coursebook on translation*. Routledge.

[9] Eddahibi, M., & Mensouri, M. (2023). English to Arabic machine translation of mathematical documents. *International Journal of Artificial Intelligence & Applications*, *14*(6). https://doi.org/10.5121/ijaia.2023.14602

[10] El-Farahaty, H., & Alwazna, R. (2024). Translating Cultural References and Sensitive Elements Between Arabic and English: Netflix Subtitles as a Case Study: *SAGE Open*, *14*(2), 1-15. DOI: 10.1177/21582440241258293

[11] Faiq, S. (2019). Discourse in Arabic translation. In K. Washbourne & B. Van Wyke (Eds), *The Routledge Handbook of Literary Translation* (pp. 433-447). Routledge.

[12] Fayed, W. (2007). *A dictionary of collocations in contemporary Arabic: Arabic-Arabic* [in Arabic]. Cairo University Press.

[13] Hafiz, al. T. (2004). *Al-Hafiz Arabic Collocations Dictionary*. Maktabat Lobnan Nashiroon.

[14] Hao, R., Hu, L., Qi, W., Wua, Q., Zhang, Y., & Nie, N. (2025). ChatLLM network: More brains, more intelligence. *AI Open*, *6*, 45–52. https://doi.org/10.1016/j.aiopen.2025.01.001

[15] Hassan, H., Aue, A., Chen, C., Chowdhary, V., & Wu, H. (2018). *Achieving human parity on automatic Chinese-English news translation*. https://doi.org/10.48550/arxiv.1803.05567

[16] Heliel, M. (2000). In the making: A new lexicon for translation from Arabic into English (in Arabic). *Alam Alfikr*, *28*(3), 223-249.

[17] Jiao, W., Wang, W., Huang, J. T., Wang, X. & Tu, Z. P. (2023). *Is ChatGPT a Good Translator? Yes with GPT-4 as the Engine*. https://doi.org/10.48550/arXiv.2301.08745

[18] Koehn, P., & Knowles, R. (2017). Six challenges for neural machine translation. In T. Luong, Birch, A., Neubig, G., & Finch, A. (Eds.), *Proceedings of the First Workshop on Neural Machine Translation* (pp. 28-39). Association for Computational Linguistics. DOI: 10.18653/v1/W17-3204

[19] Munday, J. (2016). *Introducing translation studies* (4th ed.). Routledge.

[20] Pedersen, J. (2011). *Subtitling norms for television: An exploration focusing on extralinguistic cultural references*. John Benjamins.

Said Faiq is a professor of intercultural studies and translation at the American University of Sharjah. Prior to this, he worked in Africa, the Middle East and the United Kingdom. He is a teacher and researcher of inter/cultural studies. His research sits at the interface of intercultural communication, media and representation, translation/interpreting, English and linguistics, and aspects of innovation/entrepreneurship, and AI applications to communication. orcid.org/0000-0003-3467-3791. Email: sfaiq@aus.edu