Research Journal in Advanced Humanities

RJAH

https://doi.org/10.58256/k4d6pp20

Open Access

Check for updates

RESEARCH ARTICLE **Section:** *Literature, Linguistics & Criticism*

A reception study of AI-translated idioms and proverbs between Arabic and English

Noor Darwish[1], Ahmad S Haider[1], Bayan Tannous[1], Ronza N. Abu Rumman[2], Doaa Alantari[3], Hadeel Saed[1] & Mohammed Dagamseh[4]

[1]Department of English Language and Translation, Applied Science Private University, Amman, Jordan

[2]Department of English Language and Literature, University of Jordan, Amman, Jordan

[3]Ministry of Education, Jordan

[4]Department of Linguistics, University of Canterbury, Christchurch 8140, New Zealand

*Correspondence: a_haidar@asu.edu.jo

ABSTRACT

This study investigates the perceptions of students toward the English-Arabic AI-translation of idioms and proverbs. This involves the evaluation of the performance of ChatGPT when translating Jordanian idiomatic expressions and proverbs into English, using a dataset of 100 examples. Drawing on a large-scale sample of 3,000 students from public and private Jordanian universities, the research examines how students engage with AI tools in translation of culturally embedded figurative language. Using a structured 25-item questionnaire, the analysis revealed that participants who frequently encountered idiomatic content and received formal training in translation evaluation demonstrated greater scepticism toward AI's handling of figurative meaning. Furthermore, while students acknowledged the efficiency and utility of AI tools, a strong preference remained for human translation when dealing with nuanced, culturally rich expressions. The assessment of GPT's performance in rendering the idioms showed inconsistency with most renditions relying on literal translation. Figurative renditions and paraphrases were significantly less common. This implies moderate or limited linguistic ability and cultural sensitivity. These findings can benefit AI developers, linguists, and subtitlers by showcasing some current shortcomings in AI translation.

KEYWORDS: AI translation, idioms, proverbs, Arabic-English translation, pedagogy, figurative language, human vs. machine translation, student perceptions

Research Journal in Advanced Humanities

Volume 6, Issue 3, 2025

ISSN: 2708-5945 (Print)

ISSN: 2708-5953 (Online)

ARTICLE HISTORY

Submitted: **08 June 2025**

Accepted: **12 August 2025**

Published: 25 August **2025**

HOW TO CITE

Darwish, N., Haider, A., Tannous, B., Rumman, R. N. A., Alantari, D., Saed, H., & Dagamseh, M. (2025). A reception study of AI-translated idioms and proverbs between Arabic and English. *Research Journal in Advanced Humanities*, *6*(3). https://doi.org/10.58256/k4d6pp20

Published in Nairobi, Kenya by Royallite Global, an imprint of Royallite Publishers Limited

Introduction

This study draws on recent scholarship addressing semantics, pragmatics, education, and technology in language and translation studies (Al-Dhuhli, Alkindi, & Al-Taani, 2022; Ammari & Al-Ahmad, 2023; Fraihat, 2024; Ja'afreh, 2023; Khatatbah & Ashour, 2024; Masoud, 2025; Meqdad, Al-Bayyari, & Al-Taher, 2023; Naib, 2025). In recent years, AI-powered translation tools have gained widespread recognition due to their ability to translate texts quickly and somewhat accurately. Zong (2018) claimed that machine translation (MT) outperforms human translators in terms of rapid adaptation to new terminology and bespoke lexicon, with its capacity to translate tens of millions of documents daily. Adelnia and Dastjerdi (2011) highlighted that idioms are lexical expressions of language that relate to a culture and refer to ideas, or phenomena of life and are valuable to any language in maintaining its regional and cultural specificity. In addition to that, idiomatic expressions remain significant challenges for accurate translation. According to Larson (1984), "an idiom is a string of words whose meaning is different from the meanings of the individual words" (p. 20).

Idioms, with their metaphorical and cultural nuances that exceed the literal meaning of words, cause a particular issue for MT. Hasna'a (2023) supported this notion by arguing that language development in these systems cannot follow through without human intervention. To build upon that, this study investigates the ability of AI-powered systems to translate idiomatic expressions and proverbs from Arabic into English. This paper examines the AI tool ChatGPT 3.5 (GPT), which could be used as a translation tool. GPT is an AI chatbot developed by the company OpenAI and is designed as an advanced conversational AI that is intended to work closely with user instructions and provide complete responses. According to OpenAI, GPT's conversational format makes it one of the few AI tools capable of being used to follow-up conversation questions, self-correct for mistakes, identify and push back on faulty premises, and deny inappropriate requests. GPT utilizes numerous natural language processing capabilities, such as question answering, storytelling, logical reasoning, and machine translation (Jiao et al., 2023).

Several scholars have studied the effectiveness and quality of GPT in terms of translations. As for whether GPT can translate Arabic proverbs and idiomatic expressions to English, further research is still needed. Results may vary based on the valuation criteria such as accuracy and fluency. This research, by analyzing the responses of 3000 university student to the performance of AI tools in rendering figurative language, and whether the translation is natural and accurate, will likely offer useful benefits for translators and an AI developer. The former will have a better understanding of AI tools, ChatGPT in this case, and assist latter in judging the AI systems' reliability when handling culture-bound expressions. Translators and AI developers will probably be able to reflect on the types of errors ChatGPT creates, so that they can improve existing and future language processing models.

This study attempts to answer the following research questions:

1. How do translation students evaluate the degree of accuracy and naturalness of AI-generated translations of idioms and proverbs between Arabic and English?
2. What are the main approaches followed by ChatGPT when translating Arabic idiomatic expressions and proverbs into English?

Literature Review

Idiomatic expressions are characterized by their special features. Mariana et al. (2021) define idioms as phrases whose meanings differ from those of their components, thus the meaning of the words as one unit is different from their individual meanings. Meryem (2010) and Zayed et al. (2021) suggested that their unique qualities make them different from the standard language. They added that idioms and frozen expressions operate as a single entity, including a fixed grammatical structure along with a fixed word order that cannot be changed. In light of this, translating them is a hard task due to the above mentioned reasons. Additionally, Bahumaid (2010) argued that they are not only associated with informal spoken language but are also common in written forms, such as stories and articles in magazines and journals.

Calioglo and Sandu (2023) stated that phraseological units such as idioms are reflections of culture present in everyday speech. On a similar basis, Al-Shawi and Mahadi (2012) argued that translating idiomatic expressions, especially for non-literal counterparts, requires having enough knowledge of each language's

culture. This is supported by Ali and Al-Rushaidi (2016), explaining that such culturally bound elements have long been a source of difficulty because of their deep roots in specific cultures. These expressions provide a detailed picture of people's traditional means of experiencing reality in addition to their "values and warnings, and rules and wisdom the elders want to impress on the minds of their young" (Al-Shawi & Mahadi, 2012, p. 140).

Azizov (2024) argued that machine translation and AI systems may have difficulties in the rendition of idioms as they are built on social, historical, and cultural contexts. Following this, Al-Shawi and Mahadi (2012) showed the limitations of literal translation, in which "an idiom cannot be understood simply from the meaning of its individual words, i.e., it is not total of the meaning of its individual parts which pose a serious problem for the translator" (p. 140). By way of illustration, Al-Shawi and Mahadi (2012) and Bahumaid (2010) underscored the failures of employing literal translation when dealing with idioms, even though translators use this typical approach to attain accuracy and faithfulness.

Khoshafah (2023) examined the ability of ChatGPT 3.5 to translate texts from different genres between English and Arabic accurately by comparing ChatGPT's translations to human translations. The findings showed that ChatGPT can translate simple contexts, which eventually need human intervention. It ensures the need for awareness when using this tool in translating such texts, in addition to translation assessment and revision by humans, ensuring that the translated piece is accurate enough.

Jiao et al. (2023) provided an evaluation of ChatGPT as a translation tool by testing different translation prompts to rate their performance and found that all of them produced good translations. The same study compared ChatGPT to commercial translation systems, such as Google Translate. Regarding accuracy, the outcomes showed that ChatGPT is better at translating high-resource languages but not that well in the case of low-resource languages. In addition, ChatGPT performed well in spoken language, more than in biomedical abstracts or Reddit comments. The study also explored a new technique called "pivot prompting" to solve the problem of translating low-resource languages. This is achieved by asking ChatGPT to convert the provided text into a high-resource language and then translate it into the required language. The study concludes that the ChatGPT-4 engine enhances the quality of ChatGPT's translation even for low-resource languages.

Rusadi and Setiajid (2023) investigated the performance of Google Translate and ChatGPT in translating the GUI texts of Windows 11 into Indonesian. The goal of determining the type of these errors is to assist the developers in finding a solution and fixing the algorithms in the future. The results showed various errors that were classified based on Koponen's approach.

On the other hand, Yilmaz, Naumovska, and Aggarwal (2023) discussed the threat posed by AI tools in taking human jobs, particularly translators. The study examined the neural network-based translation system and ChatGPT's translation ability. The results found that these tools reduced the number of translation jobs requiring analytical rather than emotional or cultural translation. Aldelaa (2024) shed light on the problems that AI systems have when translating Arabic idioms into English. It examined some samples from Arabic idioms extracted from three books, "The Season of Migration to the North," "The Fall of the Imam," and "The Girls of Riyadh," and compared them to both the English translations produced by MT tools and provided human translations. It uncovered the special issues that arise when translating Arabic-English idioms.

Lastly, Al-Salman and Haider (2024) evaluated the quality of popular MT and AI systems, namely, Google Translate, ChatGPT, and Gemini when translating research titles in humanities and social sciences from Arabic into English. This study attempts to help researchers identify the best system for translating their research titles. As a result, they found that the three systems achieved different errors regarding the sense or sentence structure. It is worth noting that, while human translations also included mistakes, particularly in grammar, they had the most accurate vocabulary. Gemini leads the curve in this domain among the three systems since it produced the lowest number of errors.

As proved by previous research, despite AI tools showed noticeable progress in translation, they still encounter several issues. These include failing to fully understand the intended meaning. In addition, some systems tended to employ word-for-word translation, which caused various difficulties regarding language complexities such as lexical ambiguity. As a result, the studies proved that human intervention is critical in improving MT when dealing with fixed expressions and the context-misleading nature of spoken language.

Although several studies examined the efficiency of Machine translators in dealing with idiomatic

expressions and proverbs from Arabic into English and vice versa, little attention has been paid to how users evaluate the quality of translation produced by AI tools such as ChatGPT. Therefore, this study fills this gap by examining the perceptions of 3000 translation students toward AI-translated idioms and proverbs between Arabic and English.

Method

Quantitative Part

In recent years, Jordan has witnessed significant growth in online technology adoption (Al-Salman & Haider, 2021; Al-Salman, Haider, & Saed, 2022; Haider & Al-Salman, 2020). The study involved 3,000 translation students from public and private Jordanian universities during the 2024–2025 academic year, representing all academic levels from undergraduate to postgraduate. A structured questionnaire was used to assess students' perceptions, experiences, and preferences regarding AI-translated idioms and proverbs, comprising two sections: demographic information and 25 items across six thematic dimensions.

Reliability analysis using Cronbach's alpha was conducted for each thematic construct, with all values falling within acceptable reliability thresholds (ranging between 0.73 and 0.82), as Table 1 shows.

Table 1. Internal Consistency (Cronbach's Alpha) for Each Construct

Construct	Items	Cronbach's α
Exposure and Confidence in Translating Idioms and Proverbs	1–4	0.75
Use and Frequency of AI Translation Tools	5–8	0.78
Perceived Accuracy and Error Detection in AI-Translated Idioms	9–12	0.76
Attitudes Toward the Role of AI in Translation	13–16	0.74
Preference for Human vs. AI Translation in Idiomatic Contexts	17–21	0.79
Willingness to Use AI Tools for Academic/Professional Translation	22–25	0.82

Regarding ethical approval, this study followed the ethical guidelines set forth by the Declaration of Helsinki. Ethical approval was granted by the Institutional Review Board at The Applied Science Private University in Jordan (Approval Code: FOAH 3/2024, dated 28/02/2025).

Translation Assessment

Corpus Size

The corpus of this study consisted of a collection of 100 Jordanian idiomatic expressions and proverbs and their English renditions. The expressions were extracted from nine websites, namely learnenglish. nu, wisdomarabic. com, lawjo.net, io.hsoub.com, sqorebda3.com, trend.nl7za.com, kulalsalafiyeen.com, quizlet.com, and eqrae. com. For the expressions to be chosen, they had to be commonly used by native Arabic speakers in daily conversations. After the expressions were chosen, they were translated into English by GPT. The AI tool was given the prompt, "Translate the following Arabic idiomatic expressions and proverbs into English".

Investigated AI Tool (ChatGPT 3.5)

Işım and Balcioglu (2023) highlighted that ChatGPT is an OpenAI-developed generative pre-trained transformer model that appeared in 2022. Jiao et al. (2023) showcased the range of tasks GPT can handle, including question-answering, writing, summarizing, translation across languages, storytelling, and even code debugging.

Data Analysis Approaches

The framework adopted in this study is structured as follows:

1) **Literal translation:** this approach aims to keep the ST's structure and words without focusing on the intended message.
2) **Paraphrasing the sense with common language:** this approach captures the overall idea by changing the ST's structure.
3) **Translation through figurative language:** this is achieved by translating the idiom into an idiom or an expression that includes figurative language.

Study Procedures

The procedures and steps followed in this study are:

- Selecting the translation tool (namely GPT was selected).
- Selecting the type of data to be investigated (Arabic idiomatic expressions and proverbs).
- Selecting different websites that upload Arabic idiomatic expressions and proverbs along with their English counterparts: learnenglish.nu, wisdomarabic.com, lawjo.net, io.hsoub.com, sqorebda3.com, trend.nl7za.com, kulalsalafiyeen.com, quizlet.com, and eqrae.com.
- Extracting 100 Arabic common idiomatic expressions and proverbs and their English counterparts from the selected websites.
- Copy-pasting the Arabic idiomatic expressions and proverbs to the system to be translated into English. The following prompt was used: "*Translate the following Arabic idiomatic expressions and proverbs into English*".
- Evaluating the results in terms of *literal translation, sense-based translation, and translation through figurative language*.
- Linking the results to the previous literature.

Results and Discussion

This section presents the study's findings. The results are organised into three subsections: (1) analysing demographic data, (2) descriptive analysis of the questionnaire items , and (3) translation assessment.

Sample Characteristics

Demographic profile of the participants is discussed in Table 2.

Table 2. Demographic Summary of Participants (N = 3,000)

Variable	Category	Count	Percentage (%)
Gender	Male	1,290	43.00%
	Female	1,620	54.00%
	Prefer not to say	90	3.00%
University Type	Public	1,620	54.00%
	Private	1,380	46.00%
Academic Level	Undergraduate – Year 1	570	19.00%
	Undergraduate – Year 2	600	20.00%
	Undergraduate – Year 3	660	22.00%
	Undergraduate – Year 4	690	23.00%
	Postgraduate (Master's or PhD)	480	16.00%
AI Tool Use Frequency	Daily	720	24.00%
	Weekly	840	28.00%
	Occasionally	780	26.00%
	Rarely	390	13.00%
	Never	270	9.00%
Formal Training in Translation Evaluation	Yes	1,380	46.00%
	No	1,620	54.00%
GPA Range	Excellent (3.5–4.0)	1,080	36.00%
	Very Good (3.0–3.49)	1,050	35.00%
	Good (2.5–2.99)	600	20.00%
	Fair or Below	270	9.00%
Encountering Idioms	Very often	1,140	38.00%
	Sometimes	1,020	34.00%
	Rarely	540	18.00%
	Never	300	10.00%

Table 2 indicates that there was a slight female majority (54%), which is representative of the current trend of overall enrollment in translation and other related disciplines at university level in Jordan. Types of students from both public and private universities were represented in the sample. The sample also consisted of 16% post graduates which shows the degree of involvement of senior learners in the use of AI tools for idiomatic translation. More than half of the selected students admitted to using AI translation tools at least once a day or a week, which indicates a much integrated presence of such technologies within their workflow. For example, 72% of the students indicated that they find idioms or proverbs in their translation activities "very often" or "sometimes." This puts a stronger case for the value of assessing how those expressions are treated by AI systems in addition to students' perceptions of those systems' performance in these very challenging areas of language.

Analysis of the Questionnaire's Items

Frequencies and percentages of the questionnaire items are presented in Table 3.

Table 3. Questionnaire Item Responses by Construct (N = 3,000)

Construct	Item	Questionnaire Item	Agree %	Neutral %	Disagree %
Exposure and Confidence	1	I frequently encounter culturally specific expressions such as idioms and proverbs in my translation coursework.	70.20%	18.60%	11.20%
	2	I can clearly explain the intended meaning of Arabic idioms prior to translating them.	65.80%	20.30%	13.90%
	3	I can translate Arabic idioms into accurate English equivalents.	63.70%	22.40%	13.90%
	4	I can clearly distinguish between literal and figurative meanings when translating idiomatic expressions.	68.90%	19.10%	12.00%
Use of AI Tools	5	I regularly use AI translation tools during my translation tasks.	66.50%	17.20%	16.30%
	6	I often rely on AI tools to support my understanding of complex or unfamiliar expressions.	62.10%	20.70%	17.20%
	7	I use AI translation tools to compare with or improve my own translations.	69.30%	16.90%	13.80%
	8	I frequently explore multiple AI tools to evaluate the quality of their translations.	57.80%	21.50%	20.70%
Accuracy and Error Detection	9	The meaning of idioms remains culturally appropriate after AI translation.	49.60%	25.90%	24.50%
	10	AI translations of proverbs preserve their original intent and tone.	52.30%	24.80%	22.90%
	11	I can recognize when an idiom has been translated too literally by the AI.	71.00%	15.50%	13.50%
	12	I question the reliability of AI translations when idioms or proverbs are unfamiliar to me in either language.	67.10%	18.20%	14.70%
Attitudes Toward AI	13	I believe AI plays a valuable role in improving the speed and efficiency of translation.	78.40%	13.20%	8.40%
	14	I am concerned that over-reliance on AI in translation may reduce the need for human translators.	70.70%	15.00%	14.30%
	15	I trust AI to handle everyday translation tasks, but not complex cultural expressions.	68.20%	17.80%	14.00%
	16	The integration of AI in translation enhances communication across languages and cultures.	75.10%	14.50%	10.40%

Preference: Human vs. AI	17	I prefer AI-generated idiom translations when I need a quick, general understanding of the text.	60.50%	21.60%	17.90%
	18	AI translations of idioms often feel less natural than those done by humans.	66.40%	19.70%	13.90%
	19	I prefer AI translation of idioms because it exposes me to alternative interpretations I might not expect from humans.	58.20%	24.50%	17.30%
	20	I trust human translators more because they consider the speaker's intention when translating idioms.	72.60%	15.40%	12.00%
Willingness to Use AI	21	I would use AI-generated translations of idioms and proverbs as a starting point, but rely on human review for accuracy.	76.00%	14.80%	9.20%
	22	I am willing to use AI translation tools for academic or professional tasks that do not involve complex or sensitive content.	73.10%	16.00%	10.90%
	23	I would consider using AI tools in professional translation projects to improve productivity and save time.	68.70%	18.60%	12.70%
	24	I am cautious about using AI translation tools in academic or professional contexts where precision is crucial.	64.90%	20.10%	15.00%
	25	I am open to integrating AI translation tools into my workflow if they complement human expertise rather than replace it.	79.40%	12.30%	8.30%

Over 70% of students frequently encountered idioms and proverbs, and 68.9% could distinguish between literal and figurative meanings. Overall, they showed strong confidence but still needed deeper cultural and semantic training.

AI tools were widely used, with 66.5% using them regularly and nearly 70% relying on them to refine translations. However, only 57.8% explored multiple platforms, favoring familiar tools like Google Translate and ChatGPT.

While over 70% could detect literal AI translations, only 49.6% believed AI maintained cultural meaning. Many questioned its reliability in unfamiliar contexts, showing selective trust.

Most students valued AI for speed (78.4%) and cross-cultural communication (75.1%) but remained concerned about its cultural limits and impact on human translators.

Around 60.5% used AI for quick understanding, yet 72.6% trusted humans more for interpreting intent, and 76% preferred AI only as a first draft. Students viewed AI as helpful but not a substitute for human insight.

Students showed strong openness to AI use—79.4% supported it as a complement to human expertise and over 70% used it for efficiency—though many remained cautious in tasks requiring high precision.

More than 70% of the students had quite frequent exposure to idioms and proverbs, and 66.89% could obviously distinguish between literal and figurative meanings. Nevertheless, they were quite confident and lacked some training culture-wise and even at the level of semantics.

The usage of AI was prominent, i.e., 66.5% of students used some type of AI in their studies on a regular basis, whereas almost 70% declared that they used such AI to polish translations. However, only 57.8% considered trying out more than one platform, often sticking with user-friendly Google Translate and ChatGPT. Over 70% could tell the difference between literal AI translations and those truly considered as such culturally; only about half of them felt AI was able to culturally signify anything. The rest doubted its authority on matters they were not conversant with, only believing in its decisions selectively.

As for speed, AI was accepted as an asset by 78.4% of students, while for cross-cultural communication, 75.1% saw it as being in their favor; the rest raised issues regarding cultural barriers along with its associated stigma towards human translators.

Table 4 illustrates how translations are categorized into literal and word-for-word translations, paraphrasing the sense with common language, and translation through figurative language.

Table 4: Translation processes statistics

Approach	Frequency

Literal and Word-for-Word Translations	56
Paraphrasing the Sense with Common Language	7
Translation Through Figurative Language	37
Total	100

The data used in the study comprises 100 Arabic proverbs and idiomatic expressions translated into English by GPT, with translations categorized into three strategies: Literal and Word-for-Word Translation, Paraphrasing the Sense, and Translation Through Figurative Language - with some strategies further categorized for specificity. The quantitative data suggests specific patterns in GPT's translation capabilities.

There are 56 cases identified as literal and word-for-word translations, which accounted for 56% of the dataset. This Strategy includes taking each word literally or word-for-word into the target language but rarely restructuring or reformulating the expression. Of these translations, literal makes sense is the largest subcategory, making up 32%. This represents cases in which a literal rendering happened to be both an accurate and fluent English translation. Word-for-word (confusing) accounted for 17% where the literal meaning is preserved in the target language but resulted in a confusing expression. Word-for-word (mistaken sense) occurs in 5% where literal translation results in a mistaken interpretation of the expression. Lastly, literal miscellaneous errors accounted for 2% references to errors that relate to grammar or structure preservation.

The second strategy of paraphrasing the sense in common language is very rare, accounting for only 7% of cases. This strategy emphasizes stating the meaning in plain, everyday English, rather than representing the original formal structure or figurative style. In this category, sense-for-sense translations occurred in 5%, meaning the original meaning was conveyed in clear language. sense for sense mistranslations occurred in 2%, meaning that the paraphrased version changed the meaning.

The third and second most frequently used strategy is translation through figurative language, accounting for 37%. With this method, GPT attempts to keep the figurative, metaphorical, or idiomatic nature of the source expression by choosing an equivalent or adapting a figurative expression in the target language. The largest subcategory within figurative language is idiom for the same idiom, representing 17%, where GPT translated an Arabic idiom directly into an English idiom that is accurate. Idiom for similar idiom occurred in 7%, where the translation maintained the general meaning but could be a slightly different figurative phrase. Idiom for different idioms occurred in 8%, where the idiom translated into an English idiom different from the Arabic one, achieving a slightly different meaning overall. There was a small proportion of idiom for different idioms (approximate) at 2%, idiom for different idioms (mistranslation) at 1%, and figurative language at 2%, where GPT chose metaphorical language over specific idioms.

Overall, the quantitative analysis suggests a strong propensity for literal strategies as they have accounted for more than half of all the cases, with the second category, figurative strategies, suggesting that GPT is capable of retaining idiomatic elements to some extent in a surprising proportion of translations, but not consistently yet. Furthermore, paraphrasing was extremely rare, suggesting that GPT seems to prefer relative structural fidelity or figurative modification over simply rephrasing in plain language.

Table 5 showcases some examples of GPT translations.

Table 5: Selected translations by GPT

No.	Idiomatic Expression	Literal Translation	GPT Translation
1	لذعلا فيسلا قبس	The sword preceded the blame	The sword precedes reprimand
2	ةوبك داوج لكل	Every horse stumbles	Every horse stumbles
3	على أحرّ من الجمر	On hotter than embers	On pins and needles
4	ةكلهتلا ىلإ مكيديأب اوقلت ال	Do not throw yourselves into destruction	Don't play with fire
5	ضيف من ضيغ	A drop of a larger amount	A drop in the bucket
6	لمج الو اهيف يل ةقان ال	I have neither a camel nor a she-camel in it	Neither here nor there
7	رفاك عوجلا	Hunger is a disbeliever	Hunger negates faith
8	بوقعي سفن يف ةياغل	For a purpose within Jacob's inner self	There's method to Jacob's mad-ness

9	إن كان بيتك من زجاج فلا ترمِ الناس بالحجارة	If your house is made of glass, don't throw stones at others	If your house is made of gold, do not pelt others with stones

As Table 5 shows, some idioms may sound unnatural to a native English speaker despite translating them accurately in terms of grammar and lexical choice. This is evident in GPT's translation of example 1, where "العذل" is translated as "reprimand." This proverb uses "العذل" to emphasize the futility of reprimanding and blaming oneself, as well as regretting after something has already occurred. To make matters worse, GPT used the present tense "precedes" for the past tense "سبق" (preceded), which might mislead the readers into thinking that the proverb is a general statement or fact, while it refers to a certain situation that happened and finished in the past. These examples emphasize how the sense of an idiom is separate from its wording. Thus, even an accurate translation of the phrase's components would not always result in an acceptable rendition that reflects the intended meaning.

In example 2, "every horse stumbles." is translated literally. At first glance this rendition may seem straightforward and clear enough, but this is not the case. The cultural message that gives such figurative expressions their depth is not conveyed. However, the original uses the image of a stumbling horse not to talk about horses; it rather represents humans' flaws and indicates the idea that everyone, even the skilled, might make mistakes. This example confirms that Idiom translation can sometimes be tricky, even for advanced AI systems like GPT.

GPT also employed figurative language in rendering idioms. Depending exclusively on figurative translation requires a thorough awareness of cultural background in order to avoid mistakes.

Sometimes GPT selects idioms in the TL that represent the same meaning of the Arabic idiom but appear in entirely different structures. This is yet again the result of the absence of a direct equivalent.

Examples 3 and 4 are two instances where the STs share similarities with their translations. Despite their similarity, they still appear to be different. Accordingly, 3, shows a similarity in the use of painful pictures. In contrast, 4 employs an entirely different picture. As for example 3, while the ST and TT share a slight similarity in using metaphors of painful elements to represent feelings of unease, the two differ in the exact comparison. The Arabic idiom states "على أحر من الجمر" (on hotter than coal) and its rendition states "on pins and needles". Despite the difference, both capture the same feelings of extreme anxiousness and restlessness. The Arabic expression emphasizes the scorching feeling that comes with impatient anticipation by picturing hot coals. Comparably, the English phrase expresses the same uneasy sensation of being on edge by conjuring up the image of sharp pins and needles. While considering the cultural differences in metaphors and images, the typical human sense of tense waiting is adequately communicated by both formulations.

In terms of utilizing entirely different metaphors, the Arabic "لا تلقوا بأيديكم إلى التهلكة," found in 4, is translated into "don't play with fire." While the Arabic proverb underlines the intentional nature of courting peril, the English phrase employs the picture of playing with fire, a volatile and possibly destructive force, to communicate the same cautionary message against recklessness and its repercussions.

Examples 5 and 6represent the only cases of the use of incorrect idioms in the dataset. These apparent mistakes are most likely due to two explanations. Firstly, the system might have emphasized a slight similarity between Arabic and English idioms. As a result, this leads to misleading translations that transfer a close essence but overlook the exact meaning. Secondly, the inexistence of corresponding idioms in the TL may cause the system to pick the closest option, even if it addresses unrelated features or affects the entire context.

GPT produced two instances of mistranslation. In example 5, despite that both represent a small amount of a larger amount, they significantly differ. The system mistakenly interprets the small part but misses the intended emphasis on the source's great amount. Thus, the Arabic idiom portrays something as a mere portion of an overflowing source. This suggests the small amount is a sample of what is yet to come. In contrast, "A drop in the bucket" focuses on the unimportance of a small amount compared to a larger needed amount. Thus, the idiom is used to express insignificance.

Additionally, GPT failed to capture the details of personal interest in something in 6. While both idioms acknowledge a lack of concrete involvement, they differ in personalization. For instance, the Arabic idiom represents a complete lack of personal benefits or stakes in a certain matter, while "neither here nor there" implies a more general message of being completely irrelevant. English idioms that better align with the ST

are the phrases “I have no horse in the race” and “I have no dog in the fight.” These idioms similarly represent personal stakes and interests with animals.

Example 7 applies figurative translation, though the resulting form is not a fixed idiom. It creates a more figurative structure, where GPT kept the surface concept of “disbeliever” but avoided the use of personification. Consequently, the original phrasing uses a metaphor personifying hunger and portraying it as a disbeliever human being. At the same time, GPT’s translation turned hunger into an action that exerts a powerful influence, making the starving person disbelieve. However, the intended meaning goes beyond its literal sense, in which the true message deviates from the religious concept of negating one’s faith, indicating that hunger can drive irrational actions and crazy behaviour.

Sometimes, GPT failed to capture the figurative language in the Arabic idioms accurately. Hence, this led to a loss of cultural meaning in the English versions. In examples 8 and 9, GPT employed figurative language yet misrepresented the intended sense. In these examples the renditions were altered versions of idioms where additions and modifications not only broke the idiomatic structure but led to shifts in meanings. As for example 8, it highlights the mistranslation of the Arabic idiom “بوقعي سفن يف ةياغل” as “there’s a method to Jacob’s madness.” Although both are related, indicating an unknown purpose behind evidently nonsensical action, the source idiom mentions the name “بوقعي” (Jacob) as a Quranic reference. The idiom references the prophet Jacob’s request for his children to enter different doors. This request is not clarified but instead described as a result of a desire within Jacob’s inner self. Since this desire is never explained, the term has come to describe a general placeholder to indicate that peoples’ intentions are not always known to others. This suggests that everyone, not a particular individual, might have hidden intentions behind their behaviors. However, including a specific person’s name, “Jacob,” in the English rendition breaks the fixed structure of the English idiom “there is method to someone’s madness” and personalizes it, losing the intended universality in the original version. Furthermore, while both idioms relate to outsiders not understanding a person’s actions, the English idiom relates to higher absurdity by describing the person as mad. This generalizes the sentiment to the person’s nature and not a specific situation.

Furthermore, example 9, contains the Arabic proverb “إن كان بيتك من زجاج فلا ترمِ الناس بالحجارة”. This idiom also exists in English along the lines of “Those who live in glass houses should not throw stones.” Yet, the translation strays from this near direct idiomatic translation and replaces the term “glass” with “gold.” The original version uses the term “glass,” indicating that the critics’ houses are made of fragile see-through material. Therefore, they should not risk attacking others as their houses that can easily get broken when criticizing others for certain faults they themselves have. However, GPT preferred using a metaphor, replacing the house’s weak material with “Gold.” Using “Gold” in the English translation suggests an indestructible valuable material, which completely misses the emphasis on the hypocrisy of such critics’ and the vulnerability of their houses.

Conclusion, Recommendations, and Implications

The study of 3,000 Jordanian translation students showed that while they are confident with idioms and appreciate AI’s speed and convenience, they remain cautious about its ability to convey cultural and figurative meanings accurately. Most preferred human translation for nuanced tasks, emphasizing the need for translator training that blends AI literacy with cultural awareness, critical thinking, and ethical understanding to balance technological efficiency with human insight.

This paper looked into the ways in which the AI system ChatGPT handles idiomatic expressions. The meaning of these expressions are non-literal which could pose a challenge to translators especially machine translators which do not understand cultural and linguistic nuances. The translation of idioms is often associated with three main strategies aside from omission which involve literal translation, replacing the ST idiom with a TT idiom, and paraphrasing. Machine translators do not follow the same methodic approaches to translation as human translators resulting in straying from these strategies. This manifested in the various subcategories discussed in this study.

The findings of this paper showed that GPT tended to rely heavily on literal translation and word-for-word translation (56%), This affirms findings of Khoshafah (2023) who also asserted ChatGPT was successful with simple text but uninformed regarding cultural or contextual sensitivity. The limited use of figurative translation in this paper (37%), is also in keeping with Khoshafah’s conclusion, which in this paper indicated

that while GPT translated the direct language, it generally did not cope well with capturing the cultural imagery common in Arabic idiomatic expressions and proverbs.

Jiao et al. (2023) and Işım and Balcioglu (2023), acknowledged GPT's competitive text-based performance in high-resource languages and declined accuracy in linguistically or cognitively far-off languages. This paper affirmed Arabic expressions represented one of these difficult linguistic issues, hence the lower proportion of figurative matches. Likewise, Işım and Balcioglu (2023) point to GPT's grammatical fluency and what appears to be a sense of correctness, even though there are still mistakes; likewise, as was shown in the present research, many of the literal outputs were incomplete in terms of grammar and meaning.

Additionally, the results reflect the studies of Rusadi and Setiajid (2023), as they noted similarities in mistranslation, omission and substitution errors when translating texts for technical professions. In both cases, distortions seemed to arise from literal translation strategies. The pattern noted by Yilmaz et al. (2023) also appears to endorse this pattern, where they noted that AI translation was more successful with analytical task as opposed to tasks with cultural nuance.

Moreover, Aldelaa (2024) provides the opportunity for the most direct comparison, as both address Arabic idioms. Despite concluding similar results showcasing that NMT systems typically fail to convey some figurative meaning, this paper highlights that idiom specific equivalence is not uniform, evidenced by various translations using unrelated idioms or idioms that were equivalent but only for the most part. Similarly, Al-Salman and Haider (2024) draw attention to errors related to equivalence and correctly matches their semantic meaning for NMT is also difficult.

In general, the comparison from these various studies demonstrates a common pattern: although GPT performs exceedingly well in terms of literal transfer and grammatical fluency, it is still limited in figurative transfer competence, especially for Arabic idioms and proverbs.The primary goal of this study is to assess the quality of GPT in translating Arabic proverbs and idiomatic expressions into English. The data available for analysis included expressions commonly utilized by native speakers, so the relevance of the work may be anchored in real-world language use.

The analysis was framed within the context of three prominent translation approaches. The first, Literal and Word-for-Word Translations, represents the most direct translation, without regard to changing the structure or meaning, which results in an awkward, and at times inaccurate, translation. The second, Paraphrase and Sense-for-Sense Translations, includes a translation of the overall meaning into a straightforward description with no figurative language used. The third, Translation Through Figurative Language, aims to maintain the original word imagery, metaphor, or idiomatic nature of the source text by finding equivalent figurative expressions in the TT language.

This study has limitations. Firstly, the data did not distinguish between Standard Arabic and vernacular forms of proverbs and idiomatic expressions or proverbs from idiomatic expressions in a way that could be useful for more targeted analysis in future studies. Future research could add specifications in these regards.

Acknowledgment

This research received grant no. (323/2024) from the Arab Observatory for Translation (an affiliate of ALECSO), which is supported by the Literature, Publishing & Translation Commission in Saudi Arabia.

References

Adelnia, A., & Dastjerdi, H. V. (2011). Translation of idioms: A hard task for the translator. *Theory practice in language studies, 1*(7), 879-883. doi:https://doi.org/10.4304/tpls.1.7.879-883

Al-Dhuhli, R., Alkindi, L., & Al-Taani, W. (2022). Digital Technology and its Role in Achieving Sustainable Development Requirements in Basic Education Schools in the Sultanate of Oman. *Jordan Journal of Applied Science Humanities Science Series (JJOAS), 31*(1), 43-62. doi:https://doi.org/10.35192/jjoas-h.v31i1.282

Al-Salman, S., & Haider, A. S. (2021). Jordanian University Students' Views on Emergency Online Learning during COVID-19. *Online Learning, 25*(1), 286-302. doi:https://doi.org/10.24059/olj.v25i1.2470

Al-Salman, S., & Haider, A. S. (2024). Assessing the accuracy of MT and AI tools in translating humanities or social sciences Arabic research titles into English: Evidence from Google Translate, Gemini, and ChatGPT *International Journal of Data and Network Science, 8*(4), 1-16. doi:doi: 10.5267/j.ijdns.2024.5.009

Al-Salman, S., Haider, A. S., & Saed, H. (2022). The psychological impact of COVID-19's e-learning digital tools on Jordanian university students' well-being. *The Journal of Mental Health Training, Education and Practice, 14*(4), 342–354. doi:10.1108/JMHTEP-09-2021-0106

Al-Shawi, M., & Mahadi, T. S. T. (2012). Strategies for translating idioms from Arabic into English and vice versa. *Journal of American Arabic Academy for Sciences and Technology, 3*(6), 139-147.

Aldelaa, A. S. (2024). Investigating Problems Related to the Translation of Idiomatic Expressions in the Arabic Novels Using Neural Machine Translation. *Theory Practice in Language Studies, 14*(1), 71-78. doi:https://doi.org/10.17507/tpls.1401.08

Ali, H., & Al-Rushaidi, S. (2016). Translating idiomatic expressions from English into Arabic: Difficulties and strategies. *Arab World English Journal (AWEJ), 7*(4), 15. doi:https://doi.org/10.24093/awejtls/vol1no3.2

Ammari, R., & Al-Ahmad, W. (2023). A corpus-based study of English synonyms: Small, little, tiny, and petite. *Jordan Journal of Applied Science-Humanities Series, 36*(2), 1-20. doi:https://doi.org/10.35192/jjoas-h.v36i2.542

Azizov, D. (2024). From Idioms to Algorithms: Translating Culture-Specific Expressions in AI Systems. *Iconic Research And Engineering Journals, 7*(10), 543-551.

Bahumaid, S. (2010). Strategies of translating idioms in English-Arabic dictionaries. *International Journal of Arabic-English Studies, 11*(1), 101-126. doi:https://doi.org/10.33806/ijaes2000.11.1.9

Calioglo, E., & Sandu, T. (2023). *The concept of idiomatic expression in translation.* Paper presented at the Актуальные проблемы лингвистики и лингводидактики в контексте современных подходов.

Dhillon, B. P. S. ., Sinurat, B. ., & Herman, H. . (2025). Leveraging YouTube and Instagram short videos as tools to overcome the limitations of traditional learning materials in improving foreign language undergraduates' writing proficiency. *Research Journal in Translation, Literature, Linguistics, and Education*, *1*(2), 33-46. https://doi.org/10.64120/6yc15a55

Fraihat, B. (2024). The Effect of Implementing a Remedial Plan Based on Reading Strategies on Improving the Reading Comprehension Achievement of Jordanian EFL Eleventh-Grade Students in the Vocational Stream. *Jordan Journal of Applied Science-Humanities Series, 39*(1), 18-26. doi:https://doi.org/10.35192/jjoas-h.v39i1.729

Haider, A. S., & Al-Salman, S. (2020). Dataset of Jordanian University Students' Psychological Health Impacted by Using E-learning Tools during COVID-19. *Data in Brief, 32*. doi:https://doi.org/10.1016/j.dib.2020.106104

Hasna'a, A.-H. (2023). The Future of Translation Profession In The Light Of Artificial Intelligence. *Journal of Reproducible Research, 2*(1), 157-165. doi:https://journalrrsite.com/index.php/Myjrr/article/view/46

Işım, Ç., & Balcioglu, Y. (2023). *ChatGPT: Performance of Translate*. Paper presented at the 3rd International ACHARAKA Congress on Humanities and Social Sciences, Acharaca.

Ja'afreh, N. (2023). A pragmatic contrastive analysis of apology strategies in Jordanian arabic and english language. *Jordan Journal of Applied Science-Humanities Series, 37*(2), 69-86. doi:https://doi.org/10.35192/jjoas-h.v37i2.601

Jiao, W., Wang, W., Huang, J., Wang, X., & Tu, Z. (2023). Is ChatGPT a good translator? Yes with GPT-4 as the engine. *arXiv preprint arXiv:2301.08745*. doi:https://www.researchgate.net/publication/367359399_Is_ChatGPT_A_Good_Translator_A_Preliminary_Study

Khatatbah, D., & Ashour, M. (2024). The Role of School Administration in Activating the Use of Information Technology at Schools in Irbid Governorate. *Jordan Journal of Applied Science-Humanities Series 38*(2), 121-137. doi:https://doi.org/10.35192/jjoas-h.v38i2.718

Khoshafah, F. (2023). ChatGPT for Arabic-English translation: Evaluating the accuracy. *Research Square*. doi:https://doi.org/10.21203/rs.3.rs-2814154/v2

Larson, M. L. (1984). *Meaning Based Translation*. London and New York: University Press of America Lanham, MD.

Mariana, B., Carolina, L., Valeria, A., Bautista, E. A., Silvia, K., & Lucía, A.-F. (2021). Functional anatomy of idiomatic expressions. *Brain Topography, 34*(4), 489-503.

Masoud, S. (2025). Investigating Teachers' Attitudes toward Using Assistive Technology in Teaching Students with Hearing Impairments. *Jordan Journal of Applied Science-Humanities, 43*(2), 157-171. doi:https://doi.org/10.35192/jjoas-h.v43i2.1693

Meqdad, S., Al-Bayyari, M., & Al-Taher, M. (2023). Semantic Disparity between Amr (Command) and Nahy (Negative Command) in the Noble Quran: An Interpretive/Contextual Study. *Jordan Journal of Applied Science-Humanities Series, 36*(1), 1-10. doi:https://doi.org/10.35192/jjoas-h.v36i1.516

Meryem, M. (2010). *Problems of idioms in translation, case study: First year master.* (Master's Thesis). Mentouri University- Constantine, Algeria.

Naib, A. (2025). Algerian Viewers' Attitudes toward Taboo Language: A Case Study of Mourad Oudia's Humorous YouTube Content. *Jordan Journal of Applied Science-Humanities Series, 44*(1), 167-191. doi:https://doi.org/10.35192/jjoas-h.v44i1.1830

Rusadi, A. M., & Setiajid, H. H. (2023). *Evaluating the Accuracy of Google Translate and Chatgpt in Translating Windows 11 Education Installation Gui Texts to Indonesian: An Application of Koponen's Error Category.* Paper presented at the English Language and Literature International Conference (ELLiC) Proceedings, Indonesia.

Yilmaz, E. D., Naumovska, I., & Aggarwal, V. A. (2023). *AI-Driven Labor Substitution: Evidence from Google Translate and ChatGPT*. INSEAD.

Zayed, M. H. T., Sulong, W. M. B. W., Husain, S. B., & Yahya, M. Y. B. (2021). Strategies used in Arabic–English translation of idiomatic expressions in the novel Men in the Sun. *Review of International Geographical Education, 11*(9), 2209-2219.

Zong, Z. (2018). *Research on the relations between machine translation and human translation.* Paper presented at the Journal of Physics: Conference Series.