

Artificial Intelligence or Cultural Intelligence? Investigating the Efficiency of Large Language Models (LLMs) in Idiom Creative Translation Tasks

Asma Aloui* M'Hamed Bougara University of Boumerdes (Algeria)

a.aloui@univ-boumerdes.dz



<https://orcid.org/0009-0008-1928-0539>

Article history (leave this part):

Submission date: 14-06-2025

Acceptance date: 31-03-2026

Available online: 20-04-2026

Keywords:

Artificial Intelligence,
creative translation, culture-bound references, idioms,
Large Language Models.

Funding:

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interest:

The author(s) have declared that no competing interests exist.

Cite as (leave this part):

Hanan Abufares Elkhimry; . (2024). Title. Journal of Science and Knowledge Horizons: 4(1), 283-293.

<https://doi.org/10.34118/jskp.v2i02.2727>



The authors (2026). This Open Access article is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License (CC BY-NC 4.0) (<http://creativecommons.org/licenses/by-nc/4.0/>). Non-commercial reuse, distribution, and reproduction are permitted with proper citation.

Abstract:

The translation of idioms has always been a challenging task for translators due to their culture-bound references and level of creativity. With the rise of Artificial Intelligence technologies, translators and language instructors are using AI to assist them in different language-related tasks; however, the main concern is whether these AI technologies are efficient in creative translation tasks like rendering cultural subtleties and nuances in idioms. The present paper, which falls into the scope of AI-based translation, aims at investigating the efficiency of Large Language Models in translating idioms from Arabic into English. To attain this aim, two AI models ChatGPT and Microsoft Copilot are tested in the English translation of five commonly-used Standard Arabic idioms. After receiving the same prompts, each model's output is analyzed to verify correctness and accuracy within Venuti's foreignization and domestication framework. The findings validate the efficiency of both ChatGPT and Microsoft Copilot in rendering idioms with the former opting for foreignization approach and the latter for a combination of foreignization and domestication approaches. The findings also pinpoint the potential of AI in creative translation instruction.

Asma Aloui

Résumé:

La traduction des expressions idiomatiques a toujours constitué une tâche difficile pour les traducteurs en raison de leurs références culturelles spécifiques et de leur niveau de créativité. Avec l'essor des technologies d'intelligence artificielle, les traducteurs et les enseignants en langues utilisent l'IA pour les assister dans diverses tâches linguistiques ; cependant, la principale préoccupation reste de savoir si ces technologies sont efficaces pour les tâches de traduction créative, telles que le rendu des subtilités et des nuances culturelles présentes dans les idiomes. Le présent article, qui s'inscrit dans le domaine de la traduction basée sur l'IA, vise à examiner l'efficacité des grands modèles de langage dans la traduction d'expressions idiomatiques de l'arabe vers l'anglais. Pour atteindre cet objectif, deux modèles d'IA, ChatGPT et Microsoft Copilot, sont testés dans la traduction en anglais de cinq idiomes couramment utilisés en arabe standard. Après avoir reçu les mêmes instructions, les résultats de chaque modèle sont analysés afin de vérifier leur justesse et leur précision selon le cadre de la *foreignisation* et de la *domestication* de Venuti. Les résultats valident l'efficacité de ChatGPT et de Microsoft Copilot dans la traduction des expressions idiomatiques, le premier optant pour une approche de *foreignisation* et le second pour une combinaison d'approches de *foreignisation* et de *domestication*. Les conclusions soulignent également le potentiel de l'IA dans l'enseignement de la traduction créative.

Mots-clés : Intelligence artificielle, traduction créative, références culturelles, idiomes, Grands Modèles de Langage.

*Asma Aloui

Introduction

The lack of trust in Machine Translation (MT) and Artificial Intelligence (AI)-based output is explicably fed by reasons like inaccurate interpretations, error-prone transfers, and lack of emotional intelligence namely when rendering cultural references and nuances. However, recent advances in the field of MT and AI turned skepticism into enthusiasm.

The rise of Large Language Models (LLMs) in language education and translation offers time and effort-saving alternatives to the otherwise still efficient traditional approaches. LLMs are defined as “a type of Artificial Intelligence (AI) that are designed to understand, generate, and manipulate human language at an advanced level” (Johnsen, 2024, p.70). LLMs feature a number of advantages (Johnsen, 2024) which include but are not restricted to large-scale data processing, deeper understanding of natural languages, coherent text generation, more acute context awareness, adaptability and trainability, and inclusion of multimodal analysis.

With all these advanced features; however, some mishaps can still be detected in in MT and AI-based output particularly in regards to the translation of idioms. The present paper aims at investigating the efficiency and cultural sensitivity of LLMs in rendering idioms in Arabic-English language pair.

In the following sections, the author will investigate the role of LLMs in the translation of idioms as part of creative translation tasks. A literature review with relevant studies is provided; at first, to explore the scope of this topic. Then, the author will discuss some characteristics of creative translation comprising the strategies of rendering idioms, with particular focus on automated and AI-based translations. Limitations and ethical considerations are discussed in the subsequent section followed by methodology description, results’ presentation and discussion, and a conclusion to sum up the main points discussed throughout the paper.

1. Literature Review

The following section is dedicated for the review of relevant literature about LLMs in language education and translation.

1.1. Brief History of LLMs

According to Johnsen (2024), LLMs have witnessed tremendous developments throughout the years. The early attempts in the 1950’s and 1960’s came in the

form of what is termed Rule-based Models; they were inspired by Chomsky's transformational grammar and used handcrafted rules and syntactic structures to help with language processing. The second phased (1980's-1990's) was marked by the advent of Statistical Models and a shift to data-driven language processing and probability distributions and predictions. Later on in the 2000's and 2010's, the revolutionary Neural Networks (NN) came of age offering more advanced architectures for natural language processing (NLP) with word embedding and vectors as key concepts, and Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) as prominent variants. The implementation of Attention Mechanisms in 2014 and Transformers in 2017 upgraded NLP by processing full input sequences rather than single isolated units, and by developing mechanisms entirely reliant on self-attention. Lately, the introduction of Pre-trained Language Models like BERT in 2018 marks a milestone in the history of LLMs as these models can be fine-tuned and have the ability to process large corpora bi-directionally without supervised learning. The latest advances in the field can be seen starting from 2020 in models like Generative Pre-trained Transformer (GPT) with all its variants. These models developed massive parameters and can generate human-like texts and perform other complex tasks with minimal instruction.

1.2. AI in Translation Pedagogy

Kanglang and Afzaal (2021) offer a critical overview about the use of MT and AI in the field of translation teaching. The authors acknowledge the developments brought by technology to the field of translation studies and translation practice; however, they raise concerns about the future and the role of the human instructor within the educational experience. In the same vein, the authors point out to the economic, political and ideological implications of future technologies namely in the fields of education and translation.

Gassor (2024) highlights the significant technology-motivated changes in foreign language classes; generally, and translation classes, in particular. She stresses the importance of integrating MA and AI technologies in university curricula as a method of shaping future translators' competence in a rapidly-evolving market, and calls for collaborative work between language educators and AI specialists for better integration of technology in translation and language classes.

In an attempt to practically integrate the use of technology in translation curricula design, Gassor (2024) suggests a practical detailed twenty-week

curriculum design based on four modules spanning over sixteen weeks, with the extra four weeks for practice and workshops. The first module presents key concepts in translation and translation technologies, the second module examines contemporary translation theories like discursive approaches and gender issues in translation, the third module explores translation as an interdisciplinary activity focusing on linguistic and cultural contexts, and the fourth module tackles translation from professional perspectives.

Students are to be divided on two groups undergraduates and postgraduate each performing translation tasks and machine translation editing tasks. Translations were analyzed and surveys were conducted later on to examine the effectiveness of implementing technology in both curricula design and student' competence development. The findings showed that MT post editing activities resulted in enhancing translation quality and that the most efficient MT tools were Google Translate, Reverso Context Translation and Deepl.

Similarly, Kolhar and Alameen (2021) suggest a MT system that aims to enhance the teaching and learning experiences of non-native Arabic-speaking teachers and native Arabic-speaking students in Saudi Arabia. The system, which uses AI technologies and neural machine translation (NMT) models, is said to facilitate comprehension, enrich vocabulary intake, and help students be more engaged in class.

In a recent joint international conference on computational linguistics, language resources and evaluation, Guo *et al.* (2024) proposed an innovative method for optimizing the performance of LLMs in regards to low-resource languages. The authors argued that due to lack of pre-training data, LLMs are unable to efficiently translate low-resource languages.

Based on human experience with language learning, Guo *et al.* introduced the Translate After LEarNing Textbook (Talent) approach which mimics humans structural learning from textbooks. This approach follows a 3-step procedure:

- a. Crafting a specialized textbook for each low-resource language;
- b. Absorbing content and learning syntax patterns from the textbooks;
- c. Incorporating the learned patters to improve translation quality and accuracy.

The findings from the experiment, which was tested on 112 low-resource languages, showed an increase of 14.8% in translation performance of LMM models ChatGPT and BLOOMZ based on the acquired vocabulary and syntax

patterns from the textbooks proving; hence, that low-resource languages do not need to depend on high-resource languages to perform translation tasks.

Zero-shot translation is another common issue that faces AI models when performing translation tasks. Although it might be efficient in language pairs like English and French, but the generalization of the learned patterns from high-resource language pairs to low-resource language pairs can result in off-target errors in the input. To address this concern, Zan *et al.* (2024) propose a two-stage fine-tuning algorithm that enhances translation instruction alignment by using maximum likelihood estimation (MLE) and learning correct alignment from contradicting instructions. The findings showed a considerable decrease in off-target translation errors and an overall improvement of translation quality.

Kleidermacher and Zou (2025) point out to the dominance of English language on academic research and publications which weakens global collaboration and impedes accessibility of information to non-native English speakers. The study proposes an approach which leverages LLMs to translate scientific articles from different fields into 28 languages by means of a Q&A method that uses the source text for questions and the target text for answers. Results recorded 95.9% translation accuracy with only minor adjustments needed for terminology.

The study conducted by Aleedy *et al.* (2025) sheds light on the challenges and loads of translation instruction and suggests the use of AI-powered tools to help instructors with repetitive tasks. The researchers use an English-to-Arabic parallel corpus and suggest using LLMs like GPT to transform the corpus into parallel sentences. Translation quality was assessed using cosine similarity paired with human evaluation, and the findings showed 85.2% similarity score confirming; hence, the promising potential of AI in translation education.

2. Characteristics of Creative Translation

In her article entitled Translation as creative writing practice, Fang (2021) discusses the intersection of creative writing and translation and views the latter as a form of creative writing. The author references the age-old dichotomy of faithful and free translation and argues that translation transcends the mere transfer of words to take the form of creative writing practice. Fang mentions three forms of translation: mistranslation, self-translation, and open response and maintains that they have varying degrees of authorship and could all contribute to enhancing creative writing.

Based on psychologists and neurologists views on creativity, Kußmaul (1991) suggests four stages to the creative process of translation which involve a preparatory phase, an incubation phase, an illumination phase, and an evaluation phase. This is only a clear indicator of the nexus between cognitive abilities and intuitive faculties within the process of translation (Hönig, 1989 as cited in Kußmaul, 1991).

The anthropocentricity of contemporary research, according to Alexandrovna and Borisovich (2015), imposes a multifaceted analysis of literary translations which comprises not only linguistic factors, but also non-linguistic factors including cultural, historic, and ethnic considerations. Due to its dual nature, literary translation serves; on the one hand, the source text by being faithful to it; and on the other, the target text by creating a target culture-oriented text with linguistic and cultural features that suit the target readers. This implies that literary translation is indubitably a complicated process which combines language proficiency, cultural intelligence, and creative intuition.

In the context of transmedia projects' translation, Malenova (2018) highlights the role of translators in conveying coherent narratives across different semiotic platforms (comic books, cartoon series, films, video/computer games, etc.) in a way to cater to creators and recipients. Malenova discusses the main issues encountered by translators when rendering transmedia narratives; namely, intertextuality, multimodality, and audience engagement, and suggests the following creative translation practices to resolve them:

- **Transcreation:** Transcreation involves adapting certain segments of the source text into the recipient culture in a way to maintain the style of the original material while introducing target-appealing counterparts; and thus, creating balance between source and target. This strategy is used to render elements like culture-specific references and charactonyms.
- **Transadaptation:** Transadaptation is another form of transcreation, but the changes introduced to the target version are more radical. Malenova (2018) clarifies that “By using the strategy of transadaptation the translator changes verbal, audial, and visual codes to incorporate them into transmedia project’s environment” (p.783). These changes are tailored to meet the expectations of the target audiences.
- **Transculturation:** Transculturation is the most drastic of the three strategies as it implies the complete immersion of the source production into the target culture leading; sometimes, to a distinct version of the source material. In the words of Malenova “... by using this strategy, we

can construct a new galaxy in a transmedia universe in conformity with the laws of the target culture” (p. 784).

3. Theoretical Framework

Foreignization and domestication concepts are used in this study as a theoretical framework to examine and analyze the output of the selected AI models. According to Venuti (1995), the translator using the foreignization approach produces a source-oriented text that preserves its linguistic and cultural features. In doing so, the translator accentuates the feeling of foreignness to the target reader. On the other hand, in the domestication approach, the translator produces a target-oriented text that prioritizes fluency and naturalness and is adapted to the recipient culture’s norms.

4. Translating Idioms as a Creative Task

Translating idioms is deemed as one of the most daunting tasks for translators as it requires a combination of polished translation skills, a deep understanding of both source and target cultures, and a high level of creativity and heuristic adaptation in cases of untranslatability. Idioms are associated with natural usage of language, and are defined as a group of words whose meaning is not understood from the meaning of the combined components (Newmark, 1998). *‘Every cloud has a silver lining’* is an example of an English idiom that is used to express optimism and belief that there is always a positive outcome even in negative situations. The meaning of this idiom cannot be understood from the combination of the denotative meanings of the words ‘cloud’, ‘silver’, and ‘lining’, but is rather understood in a connotative way. Furthermore, idioms are naturally used by native speakers on regular basis in daily conversations and are deeply rooted in their cultures. For this reason, rendering them from one language/culture into another is a challenging task for translators, particularly if the language pair and culture pair have very distant linguistic and cultural backgrounds.

Accordingly, idiomatic translation (Newmark, 1988) is a form of re-creative or re-productive translation as it tends to deploy idiomatic language such as colloquial expressions and idioms in the target text even if they were not used in the source text.

4.1. Strategies for Translating Idioms

In her renowned book *In other words. A coursebook on translation*, Baker (1992) suggests a set of four strategies for rendering idioms:

4.1.1. Using an Idiom of Similar Meaning and Form

In this strategy, both SL expression and TL expression share the same meaning and the same words. The idiom ‘to shed crocodile tears’ is used to describe someone making an insincere display of emotions. Similarly, the Arabic idiom ‘يذرف دموع التماسيح’ (yadhrif dumū‘ al-tamāsīḥ) has not only the same meaning, but also the same words as the English idiom:

to shed	يذرف (yadhrif)
crocodile	تماسيح (tamāsīḥ)
Tears	دموع (dumū)

4.1.2. Using an Idiom of Similar meaning but Dissimilar Form

The second strategy implies the SL and the TL expressing the same idea, but through different words. The English idiom ‘through thin and thick’ and the Arabic idiom ‘في السراء والضراء’ (fī as-sarrā’ wā ad-ḍarrā’) express the idea of giving support to someone under any circumstances, be them positive or negative, but each idiom uses different words: the adjectives ‘thin’ and ‘thick’ in English are substituted with the nouns ‘السراء’ (as-sarrā’ meaning in good times) and ‘الضراء’ (ad-ḍarrā’ meaning in bad times) all while preserving the same meaning.

4.1.3. Translation by Paraphrase

If no match can be found in the target language, translation by paraphrase, as the name indicates, renders the source idiom by paraphrase and not by an idiom, but still preserves some stylistic effects. The English idiom ‘to have no backbone’ is translated by paraphrase into Arabic as ‘ضعيف الشخصية’ (da‘īf ash-shakhṣīyah) which means a person with weak character.

4.1.4. Translation by Omission

When the meaning of the source idiom cannot be idiomatically matched or paraphrased in the target language for stylistic reasons, the fourth strategy is used to render idioms by means of omission. The Arabic idiom ‘دب الشقاق’ (dubba ash-shiqāq which means a dispute or an argument has ignited) is

rendered in English by a simple word ‘to dispute’ or ‘to argue’. Here, there is no close match for the Arabic idiom in English and the stylistic and idiomatic effects are omitted.

The following table illustrates the four strategies compared in terms of form, meaning, and idiomatic and stylistic effects.

	Form	Meaning	Idiomatic & stylistic effects
Strategy 1	+	+	+
Strategy 2	-	+	+
Strategy 3	-	+	+/-
Strategy 4	-	+	-

Table 1: Strategies for translating idioms

4.2. Automated and AI-based Idiom Translation

In terms of approaches, MT adopts initially two prominent approaches: the rule-based approach which functions on the basis of memorized rules followed by the computer to perform different translation tasks, and the corpus-based approach which uses available corpora of word-based and phrase-based models and chooses the best translation based on the highest probability. Benyahia (2024).

In recent years, the automated translation of colloquial and idiomatic expressions has been a primary focus in the field of MT. The main issue is that the so-far automated systems tend to output a word-for-word translation and disregard the fact that the meaning of an idiom is not derived from its individual components. As a result, several systems have been designed and developed to enhance the quality and accuracy of MT and NMT output. Focusing on Example-based Machine Translation, Anastasiou (2010) demonstrates how to leverage the hybrid EBMT system METIS-II to produce correct and accurate translations using a corpus and morphosyntactic rules.

Baziotis *et al.* (2022) introduces a metric called LitTER which has the ability of detecting literal translation errors without human intervention using bilingual dictionaries and reference translations increasing ultimately the accuracy of idiom translation. However, some of its limitations include the inability to deal with compound morphology and the need to be paired with other evaluation metrics for better results.

A more novel study by Benyahia (2024) explores the efficiency of AI-based online translation services in rendering idioms used in American media. The study points to the difficulty of translating idioms due to their fixed nature and culture-bound connotations. Three common online translation engines were examined in this paper: Google Translate, DeepL Translator, and Bing Microsoft Translator, with the latter proven to be the most accurate of the three scoring 68.7% accuracy level.

5. Challenges and Ethical Considerations of Using Automated and AI-based Translation

Despite the perceptible advances in the field of MT in general and AI-based translation in particular, certain issues and concerns are yet to be addressed by future research and application. Ambiguity, tone, and intertextuality come at the forefront of these issues as the automated translation systems still lack emotional and cultural intelligence compared to human translators who have the natural faculty of understanding the nuances of natural languages.

At the other end of the spectrum, some ethical concerns arise regarding the use of AI in education and translation practices, these include:

- The fear of AI replacing humans in the work place;
- Over-reliance or total dependence on technology;
- Loss of originality, creativity, and critical thinking skills;
- Subjectivity and lack of transparency.

6. Method

As stated above, the aim of this study is to explore the efficiency of LLMs in translating idioms from Arabic to English. Two popular AI models were selected for the purpose of this study to translate five common idioms in Standard Arabic into English. The models are: Microsoft Copilot and ChatGPT. Using foreignization and domestication approaches, the author will analyze the output of these models and test their efficiency, correctness and accuracy in translating the Arabic idioms. The same prompts are given to each model, then the output will be analyzed in separate tables for clearer display.

7. Results

Five Standard Arabic idioms were chosen to be used as the corpus of analysis in this study. These idioms are commonly used by native Arabic speakers and are intrinsically bound to Arabic language and Arab culture.

The following table illustrates the output of **ChatGPT**. The instruction stipulates “Translate the meaning of the following Arabic idiom: ‘idiom’ then provide an English equivalent idiom for it”. The output is divided into three columns: literal translation, meaning interpretation and equivalent English idiom(s) for clearer display of data.

SA Idiom	Literal translation	Meaning interpretation	English idiom(s)
<p>رُبَّ ضَارَّةٍ نَافِعَةٍ</p> <p>rubba ḍarratin nāfi‘atun</p>	<p>Perhaps something harmful is (actually) beneficial.</p>	<p>It expresses the idea that what seems bad or harmful at first may turn out to be good or beneficial in the end.</p>	<p>-Every cloud has a silver lining</p> <p>-A blessing in disguise.</p>
<p>كُلُّ يَغْنَى عَلَى لَيْلَاهُ</p> <p>kullun yughannī ‘alā laylāh</p>	<p>Everyone sings about his own Layla</p> <p>“Layla” here refers to a personal story, concern, or interest.</p>	<p>This idiom describes a situation where everyone is focused on their own interests, concerns, or perspectives—often ignoring others or the bigger picture. It can also imply a lack of harmony or unity in a group.</p>	<p>-Everyone is dancing to their own tune.</p> <p>-Every man for himself.</p> <p>-Each to their own.</p> <p>-They're all singing from different hymn sheets. (<i>especially when emphasizing disunity</i>)</p>
<p>رَجَعَ بِخُفَّيْ حُنَيْنٍ</p> <p>raja‘ bi khuffay ḥunayn</p>	<p>He returned with Hunayn’s sandals.</p>	<p>This idiom describes someone who returns empty-handed after failing to achieve their goal or mission—often after great effort or high expectations. It can also imply disappointment or loss.</p>	<p>-He came back empty-handed.</p> <p>-A wild goose chase. (<i>if the failure involved chasing something pointless</i>)</p> <p>-All for nothing.</p> <p>-Left high and dry. (<i>especially when abandoned or without help</i>)</p>

<p>بلغ السيل الزبي</p> <p>balagha as-saylu az-zubá</p>	<p>The flood has reached the high ground.</p>	<p>This idiom is used when a situation has gone too far or reached a breaking point—when patience runs out and action becomes necessary. It conveys that things have escalated beyond tolerable limits.</p>	<p>-The last straw. -Enough is enough. -Things have gone too far. -At the breaking point.</p>
<p>اختلف الحابل بالتابل</p> <p>ikhtalaṭa alḥābil bi an- nābil</p>	<p>The warrior's rope got mixed with the archer's rope.</p>	<p>This idiom is used to describe a situation of confusion or chaos where things or people become indistinguishable. It refers to a lack of clarity, where everything is mixed up and it's hard to tell what's what or who's who.</p>	<p>-The lines are blurred. -It's all mixed up. -All hell broke loose. (<i>if the chaos is intense</i>) -Thrown into confusion. -A complete mess. -Can't tell heads from tails.</p>

Table 2: ChatGPT output

The following table illustrates the output of **Microsoft Copilot**. The instruction stipulates “Translate the meaning of the following Arabic idiom: ‘idiom’ then provide an English equivalent idiom for it”. The output is divided into three columns: literal translation, meaning interpretation and equivalent English idiom(s) for clearer display of data.

SA Idiom	Literal translation	Meaning interpretation	English idiom(s)
<p>رُبَّ ضَارَّةٍ نَافِعَةٍ</p> <p>rubba ḍārratin</p>	<p>-Sometimes, harm can bring benefits</p> <p>-A misfortune can turn into a</p>	<p>It conveys the idea that an event that seems negative at first might lead to unexpected positive</p>	<p>Every cloud has a silver lining</p>

nāfi‘atun	blessing.	outcomes.	
<p>كُلُّ يَغْنَى عَلَى لَيْلَا</p> <p>kullun yughannī ‘alā laylāh</p>	<p>-Everyone sings about their own Layla</p> <p>-Each person is focused on their own concerns</p>	<p>It conveys the idea that people tend to prioritize their own interests, worries, or desires, often without considering others.</p>	<p>Every man for himself.</p> <p>*The phrase originates from the famous love story of Qays and Layla, where poets would use "Layla" as a symbolic name for their beloved, making it a metaphor for personal preoccupations</p>
<p>رَجَعَ بِخُفَّيْ حُنَيْنٍ</p> <p>raja‘ bi khuffay hunayn</p>	<p>He returned with Hunayn’s sandals</p>	<p>It describes someone who failed to achieve their goal and returned with nothing.</p>	<p>To go back empty-handed.</p> <p>* The origin of this idiom comes from an old Arabic tale about a shoemaker named Hunayn, who, after being insulted by a customer, tricked him into losing his camel and returning home with only a pair of shoes—thus, the phrase became a metaphor for failure.</p>
<p>بَلَغَ السَّيْلُ الرُّبَى</p>	<p>- The flood has reached the highest point</p> <p>- Things have gone too far</p>	<p>It conveys the idea that a situation has escalated beyond control, reaching a breaking point where patience or tolerance</p>	<p>The last straw.</p> <p>*The phrase originates from an old Arabic story where hunters would dig deep pits (زبى)</p>

<p>balagha as-saylu az-zubá</p>		<p>is no longer possible.</p>	<p>(az-zubá) on elevated ground to trap lions. These pits were meant to remain dry, but if a flood reached them, it signified an extreme and unexpected event; hence, the metaphor for reaching a critical threshold.</p>
<p>اختلط الحابل بالتابل ikhtalaṭa alḥābil bi an- nābil</p>	<p>Things have become chaotic and unclear</p>	<p>It describes a situation where confusion reigns, making it difficult to distinguish between different elements or people.</p>	<p>All hell broke loose. * The origin of the Arabic phrase is debated, but one explanation comes from warfare: "الحابل" (alḥābil) refers to soldiers who used ropes to capture enemies, while "التابل" (an-nābil) refers to archers who used arrows. In the heat of battle, dust and chaos would make it impossible to tell them apart, leading to the phrase's meaning of confusion. Another interpretation comes from livestock herding, where "الحابل" (alḥābil) refers to goats ready for breeding, and "التابل" (an-nābil)</p>

			refers to those that are not—if they got mixed up, it would cause problems for the herders.
--	--	--	---

Table 3: Microsoft Copilot output

8. Discussion

Idiom 1: رَبِّ ضَارَّةٍ نَافِعَةٍ (rubba ḍārratin nāfi‘atun)

ChatGPT’s literal translation of the first idiom was a simple and straightforward translation that preserved the form and style of the Arabic idiom (the Arabic adjectives ضارة (ḍārra) and نافعة (nāfi‘a) were rendered by English adjectives harmful and beneficial), while the literal translation of Microsoft Copilot was more communicative as it changed the SL categories while preserving the meaning; *i.e.*, the adjective ضارة (ḍārra) was rendered by the noun harm and the adjective نافعة (nāfi‘a) was rendered by the phrase bring benefit.

As for the meaning interpretation, both ChatGPT and Microsoft Copilot explained the meaning of the Arabic idiom correctly and accurately, only using different styles of expression.

The third column, as mentioned above, is dedicated for the display of equivalent idioms in English, and it is the main focus of analysis for it will demonstrate the efficiency of the selected LLMs in providing correct, accurate, and/or creative translations of cultural references found in the Arabic idioms. ChatGPT suggested two English idioms ‘Every cloud has a silver lining’ and ‘A blessing in disguise’ offering hence more choices, while Microsoft Copilot suggested one equivalent English idiom ‘Every cloud has a silver lining’. These English idioms convey the meaning of the source input while still deploying a target-oriented approach; the words ضارة (ḍārra) and نافعة (nāfi‘a) are inherently associated with Arab and Muslim culture as they are mentioned in religious contexts and are frequently used in moral education settings. The English idiom; on the other hand, is more familiar to Western and Christian culture and flows naturally to English language speakers; the words cloud and blessing are also of common use in the Christian community. Accordingly, both ChatGPT and Microsoft Copilot translated the first Arabic idiom using a domestication approach which assured the correct and accurate transfer of meaning and highlighted the cultural differences between Arabic and English.

Idiom 2: كُلُّ يُغْنِي عَلَى لَيْلَاهُ (kullun yughannī ‘alá laylāh)

In the literal translation of the second idiom, while ChatGPT generated a word-for-word translation in addition to a note about the noun Layla and its potential interpretations, Microsoft Copilot suggested two renderings: a word-for-word translation into English and an explanation of the entire concept of the idiom (**Each person is focused on their own concerns**).

In regards to meaning interpretation, ChatGPT gave a more inclusive interpretation of the Arabic idiom by showing that it can refer to a person being self-centered, and refer to lack or harmony and unity in a group. Microsoft Copilot’s interpretation was broader covering the general idea of the idiom. In both case, ChatGPT and Microsoft Copilot provided clear and correct meaning interpretations of the Arabic idiom.

As illustrated in the table above, ChatGPT listed four different English idioms used to convey the concept of the Arabic idiom. For instance, the first suggested English idiom ‘Everyone is dancing to their own tune’ shows figuratively how people can be focused on their own interests and concerns while ignoring others’. This idiom also depicts the Western culture’s love for dancing and music making it; therefore, a domesticated version of the Arabic idiom. This latter mentions a famous figure in the ancient Arab culture ‘Layla’ who is the beloved woman of the eminent poet Qays. Being madly in love, Qays was known for composing and reciting poems to his Layla to express his deep affection and love. Hence, the idiom became symbol of people being absorbed in their own preoccupations just like Qays was absorbed by Layla. The second suggested English idiom is also a domesticated version of the Arabic idiom because it is inspired by the Christian saying: ‘Every man for himself and God for all’. Microsoft Copilot suggested this very idiom as the one equivalent for the Arabic input. However, unlike ChatGPT, Microsoft Copilot provided a brief explanation of the emblem figures Qays and Layla to facilitate the understanding of the Arabic idiom and its cultural references, which could be considered as a foreignization approach since the target reader is brought to close contact with the source culture.

Accordingly, ChatGPT opted for a domestication approach in rendering the Arabic idiom, whereas Microsoft Copilot opted for both domestication and foreignization approaches in the process.

Idiom 3: رَجَعَ بِخُفَّيْ حُنَيْنٍ (raja‘ bi khuffay ħunayn)

Regarding the third Arabic idiom, ChatGPT and Microsoft Copilot provided the exact same literal translation by keeping the same grammatical categories (verb for verb, preposition for preposition, and noun for noun) and opting for borrowing technique to render the Arabic name حنين (ħunayn).

As far as meaning interpretation is concerned, ChatGPT offered again another more comprehensive explanation of the Arabic idiom discussing its causes (failed goal or mission) and repercussions (feeling of disappointment or loss). Microsoft Copilot; however, gave a more concise interpretation of the idiom; failing at a achieving a certain goal and returning with nothing.

Similarly to the previous idiom, ChatGPT suggested four different English idioms (He came back empty-handed; a wild goose chase; all for nothing; and left high and dry) with specifications for options 2 and 4 (review table above). One can also notice that ChatGPT listed the four choices by order of closeness to the Arabic idiom interpretation, the first being the closest to the SL input. With the same instruction as a guide, Microsoft Copilot provided one equivalent idiom (To go back empty-handed) using the verb go back instead of the verb come back. Microsoft Copilot also told the background story of Hunayn; a shoemaker who used his wits to trick a man after this latter insulted him. At the end and after losing his camel, the man returned home defeated with only a pair of shoes in hand; and hence, the story became an idiom and a metaphor for defeat and loss.

By providing only the four equivalent idioms, ChatGPT’s approach was target-oriented or; in the terms of Venuti, domesticated as it generated English language and English culture appropriate content. Besides generating a target-oriented equivalent, Microsoft Copilot supported the domesticated content with background story-telling and explanation of source cultural references; which classifies as a foreignization approach to the translation of the idiom.

Idiom 4: بَلَغَ السَّيْلُ الزُّبَى (balagha as-saylu az-zubá)

The literal translation (The flood has reached the high ground) of the fourth idiom was similar for both ChatGPT and Microsoft Copilot, the only differences being the use of the superlative form of the adjective ‘high’ in the output of Microsoft Copilot, and the use of the word ‘point’ instead of the word ‘ground’.

Microsoft Copilot also included a clarification of the literal translation to reduce ambiguity, despite not being instructed to do so.

Meaning interpretation was thoroughly discussed by both ChatGPT and Microsoft Copilot offering a clear overview of the implications of the Arabic idiom: escalated situation, reaching a breaking point, running out of patience and tolerance, and needing to take action as a result.

ChatGPT generated yet another time four choices as English equivalents (The last straw; enough is enough; things have gone too far; and at the breaking point) which offers a wider spectrum when encountering the Arabic idiom in different contexts. Microsoft Copilot generated the idiom ‘the last straw’ and complemented it with a background story explaining the references **زبي** (zubá) and the trap-setting strategy by ancient Arab hunters.

As far as the translation of the fourth idiom is concerned, one can say that ChatGPT used a domestication approach only, while Microsoft Copilot opted for a combination of both approaches for a more well-rounded translation process prioritizing target appropriation and introducing source references.

Idiom 5: اختلط الحابل بالنابل (ikhtalaṭa alḥābil bi an-nābil)

The literal translation of the fifth idiom was more literal in ChatGPT’s output than in Microsoft Copilot’s output as this later generated an interpretative sentence (Things have become chaotic and unclear) of the Arabic idiom. One mistake can be detected in the literal translation of ChatGPT. Although it made the difference between **الحابل** (alḥābil) and **النابل** (an-nābil) by translating them as warrior and archer respectively, but it used the same word ‘rope’ for both **الحابل** (alḥābil) and **النابل** (an-nābil) instead of rope for warrior and arrow for archer.

In the meaning interpretation columns, both ChatGPT and Microsoft Copilot are seen to provide clear and simple analysis of the Arabic idiom focusing on the inability to distinguish between things and people in cases of big confusion and chaos.

In the column of equivalent idioms, ChatGPT provided, not four, but six English idioms (The lines are blurred; it’s all mixed up; all hell broke loose; thrown into confusion; a complete mess; and can’t tell heads from tails) all fairly corresponding to the Arabic input with only varying degrees of language formality. This is a clear indicator that LLMs have developed an acute ability to perform complicated creative tasks and provide multiple correct and accurate

translations. In the same vein, Microsoft Copilot provided one equivalent English idiom (All hell broke loose) in addition to background story review where it explained that the Arabic idiom was inspired by two contexts. The first context is a war-related; it gives homage to the ancient Arabs' courage and skills in battlefields, and describes the image of soldiers fighting and getting mixed up in the heat of battle. The second context refers to livestock herding where goats ready for breeding get mixed up with goats that are not yet ready for breeding, resulting in chaos. This also proves that LLMs have developed cultural sensitivity and can provide deep interpretations and analysis of culture-bound references.

Concerning the translation of the fifth idiom within the foreignization and domestication context, ChatGPT used only a domestication approach by listing six English equivalents that are target text and target culture appropriate. Microsoft Copilot once more opted for both approaches since it generated a target-oriented equivalent, and provided a source-oriented review of the cultural references to make them more familiar to English readers.

Conclusion

The present study was an attempt to investigate the efficiency of LLMs and AI-based models in creative translation tasks. ChatGPT and Microsoft Copilot were tested in the translation of five common Arabic idioms within foreignization and domestication framework. After being given the same instruction, ChatGPT and Microsoft Copilot generated relatively close output. The findings suggest the following:

- When instructed to do literal translation, ChatGPT's output tended to be a word-for-word transfer in most case with minimal clarifications about cultural references, except in the case of the second idiom;
- When instructed to do literal translation, Microsoft Copilot tended to provide more interpretative literal translations and offered two interpretations in most cases;
- When instructed to explain the meaning of the idioms, both ChatGPT and Microsoft Copilot succeeded in rendering the meanings of the Arabic idioms accurately, with ChatGPT offering more inclusive interpretations and Microsoft Copilot providing more concise interpretations;

- When instructed to suggest an English equivalent for the Arabic idiom, ChatGPT provided several correct and accurate choices that are target language and target culture appropriate;
- When instructed to suggest an English equivalent for the Arabic idiom, Microsoft Copilot generated one equivalent idiom that is also target language and target culture appropriate;
- ChatGPT used a domestication approach in suggesting the English idioms by focusing solely on target language and target culture equivalence;
- Microsoft Copilot used a combination of foreignization and domestication approaches in the generation of idioms as it provided the English equivalents in addition to a background story review to elucidate the Arabic culture-bound references and make them more accessible to English readers.
- The output of ChatGPT and Microsoft Copilot represents a significant campus for translation instruction as it can facilitate and enhance the teaching/learning experience in translation classes.

The study; consequently, validates the efficiency of LLMs and AI-based models in the creative translation of idioms, with minor mishaps to be addressed. The study also recognizes the great advances brought to the field of MT and AI-based translation which resulted in more accurate transfer of cultural references and nuances.

The author recommends more extensive and inclusive research to be conducted; in the future, in fields combining Translation Studies, cultural studies and computer science.

References

- Aleedy, M., Alshihri, F., Meshoul, S., Al-Harathi, M., Alramlawi, S., Aldaihani, B., ... & Atwell, E. (2025). Designing AI-powered translation education tools: A framework for parallel sentence generation using SauLTC and LLMs. *PeerJ Computer Science*, 11, e2788. <https://doi.org/10.7717/peerj-cs.2788>
- Alexandrovna, L. E. & Borisovich, L. D. (2015). Literary translation as a creative activity. *Научный результат. Вопросы теоретической и прикладной лингвистики*, (1), 63-69.
- Anastasiou, D. (2010). *Idiom treatment experiments in machine translation*. Cambridge Scholars Publishing.
- Baker, M. (1992). *In other words. A coursebook on translation*. Routledge.
- Baziotis, C., Mathur, P., & Hasler, E. (2023). *Automatic evaluation and analysis of idioms in neural machine translation*. Presented at the Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics, Dubrovnik. Retrieved from <https://aclanthology.org/2023.eacl-main.267/>
- Benyahia, M. (2024). Examining the Efficiency of Machine Translation in Translating English Idioms used in American Media. *Journal of Translation and Language Studies*, 5(2), 43-55. <https://doi.org/10.48185/jtls.v5i2.1070>
- Fang, X. (2021). Translation as creative writing practice. *New Writing*, 18(2), 162-176. <https://doi.org/10.1080/14790726.2020.1762665>
- Gassor, D. R. M. (2024). Artificial Intelligence and Technology Implementation in Teaching Translation in Universities. *Bulletin of The Faculty of Languages & Translation*, 27(2), 351-374. <https://doi.org/10.21608/bflt.2024.399201>
- Guo, P., Ren, Y., Hu, Y., Li, Y., Zhang, J., Zhang, X., & Huang, H. (2024). *Teaching large language models to translate on low-resource languages with textbook prompting*. Presented at the Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), Torino. Retrieved from <https://aclanthology.org/2024.lrec-main.1362/>
- Johnsen, M. (2024). *Large language models (LLMs)*. Maria Johnsen.
- Kleidermacher, H. C., & Zou, J. (2025). Science Across Languages: Assessing LLM Multilingual Translation of Scientific Papers. *arXiv preprint arXiv:2502.17882*. 1-16. <https://doi.org/10.48550/arXiv.2502.17882>
- Kolhar, M., & Alameen, A. (2021). Artificial Intelligence Based Language Translation Platform. *Intelligent Automation & Soft Computing*, 28(1). <https://doi.org/10.32604/iasc.2021.014995>
- Kußmaul, P. (1991). Creativity in the translation process: Empirical approaches. In K. M. van Leuven-Zwart & T. Naaijken (Eds.), *Translation studies: The state of the art* (91-101). Rodopi.
- Liu, K., & Afzaal, M. (2021). Artificial intelligence (AI) and translation teaching: A critical perspective on the transformation of education. *International journal of educational sciences*, 33(1-3), 64-73. <https://doi.org/10.31901/24566322.2021/33.1-3.1159>

- Malenova, E. D. (2018). Creative practices in translation of transmedia projects. *Journal of Siberian Federal University Humanities & Social Sciences*, 5, 775-786. <https://doi.org/10.17516/1997-1370-0269>
- Newmark, P. (1988). *A textbook of translation*. Prentice hall.
- Newmark, P. (1998). *More paragraphs on translation*. Multilingual matters.
- Venuti, L. (1995). *The translator's invisibility. A history of translation*. Routledge.
- Zan, C., Ding, L., Shen, L., Zhen, Y., Liu, W., & Tao, D. (2024). Building accurate translation-tailored llms with language aware instruction tuning. *arXiv preprint arXiv:2403.14399*. <https://doi.org/10.48550/arXiv.2403.14399>