The North African Journal of Scientific Publishing (NAJSP)
مجلة شمال إفريقيا للنشر العلمي (NAJSP)
E-ISSN: 2959-4820
Volume 3, Issue 4, 2025
Page No: 36-48
Website: https://najsp.com/index.php/home/index
SJIFactor 2024: 5.49 **معامل التأثير العربي (AIF) 2025: 0.69** **ISI 2024: 0.696**

Evaluating the Accuracy of Machine Translation in Rendering Arabic Idioms into English: A Case Study on Google Translate and ChatGPT

Israa Bashir Elajail*
English Language Department, Faculty of Arts, Misurata University, Misurata, Libya

تقييم دقة الترجمة الآلية في ترجمة المصطلحات العربية إلى الإنجليزية: دراسة حالة على جوجل ترانسليت *Google Translate* وشات جي بي تي *ChatGPT*

اسراء بشير العجيل*
قسم اللغة الانجليزية، كلية الاداب، جامعة مصراتة، مصراتة، ليبيا

*Corresponding author: israa.b.elajail@gmail.com

Received: August 01, 2025	Accepted: November 03, 2025	Published: November 09, 2025

Abstract:
This study investigates the translation of 145 Arabic idioms by two prominent machine translation systems: Google Translate and ChatGPT. The research employs both quantitative and qualitative methodologies to analyze translation approaches and assess accuracy in conveying idiomatic meanings from Arabic to English. Data collection involved Arabic idiomatic expressions from literary sources, cultural texts, and linguistic databases. The analysis framework builds upon Baker's (1992) taxonomy of translation strategies. Quantitative findings reveal that Google Translate employed literal translation in 74% of cases, while ChatGPT demonstrated more varied approaches with 48% literal translations. For sense-based translations using non-figurative language, ChatGPT led with 41%, compared to Google Translate's 15%. When examining figurative language translations, ChatGPT achieved 11% compared to Google Translate's 11%. The qualitative analysis highlights persistent challenges in both systems regarding cultural context preservation and semantic accuracy. The study concludes that while technological advances have improved machine translation capabilities, rendering Arabic idioms into English remains problematic due to cultural-linguistic gaps and contextual complexities inherent in idiomatic expressions.

Keywords: Google Translate, ChatGPT, Arabic idioms, English translation, Machine translation, Cross-cultural communication.

الملخص
تبحث هذه الدراسة في ترجمة 145 مصطلحًا عربيًا باستخدام نظامين رائدين للترجمة الآلية: جوجل ترانسليت *Google Translate* وشات جي بي تي *ChatGPT*. يستخدم البحث منهجيات كمية ونوعية لتحليل مناهج الترجمة وتقييم دقة نقل المعاني الاصطلاحية من العربية إلى الإنجليزية. تضمن جمع البيانات تعابير اصطلاحية عربية من مصادر أدبية ونصوص ثقافية وقواعد بيانات لغوية. يعتمد إطار التحليل على تصنيف بيكر (1992) لاستراتيجيات الترجمة. تكشف النتائج الكمية أن جوجل ترانسليت استخدم الترجمة الحرفية في 74% من الحالات، بينما أظهرشات جي بي تي *ChatGPT* مناهج أكثر تنوعًا بنسبة 48% للترجمات الحرفية. أما بالنسبة للترجمات القائمة على المعنى باستخدام لغة غير مجازية، فقد تصدّر شات جي بي تي *ChatGPT* بنسبة 41%، مقارنةً بنسبة 15% لجوجل ترانسليت. عند دراسة ترجمات اللغات المجازية، حقق شات جي بي تي *ChatGPT* نسبة 11% مقارنةً بـ 11% لـجوجل ترانسليت *Google Translate*. يُبرز التحليل النوعي التحديات المستمرة في كلا النظامين فيما يتعلق بحفظ السياق الثقافي ودقة الدلالات. وخلصت الدراسة

إلى أنه على الرغم من أن التقدم التكنولوجي قد حسّن قدرات الترجمة الآلية، إلا أن ترجمة التعبيرات الاصطلاحية العربية إلى الإنجليزية لا تزال تُمثل مشكلةً نظرًا للفجوات الثقافية اللغوية والتعقيدات السياقية الكامنة في التعبيرات الاصطلاحية.

الكلمات المفتاحية: جوجل ترانسليت *Google Translate*، شات جي بي تي *ChatGPT*، التعبيرات الاصطلاحية العربية؛ الترجمة الإنجليزية، الترجمة الآلية، التواصل بين الثقافات.

Introduction

Machine translation technology has fundamentally transformed global communication patterns, particularly in bridging the complex linguistic divides between Arabic and English (Al-Salman & Haider, 2024). This technological revolution has democratized access to cross-linguistic information exchange, enabling unprecedented levels of international collaboration and cultural interaction. However, despite remarkable advances in computational linguistics and artificial intelligence, the translation of idiomatic expressions remains one of the most formidable challenges confronting automated translation systems. Arabic, with its extraordinarily rich metaphorical heritage and profound cultural specificity, presents unique and multifaceted obstacles when translating figurative language into English (Almrayat, Farghal, & Haider, 2024).

The historical development of machine translation has witnessed several paradigm shifts, from rule-based systems to statistical approaches, and most recently to neural machine translation architectures (Zakraoui, Saleh, Al-Maadeed, & Alja'am, 2021). Each evolutionary phase has brought improvements in handling straightforward lexical translations, yet idiomatic expressions continue to pose significant challenges due to their non-compositional nature and cultural embeddedness. This persistent difficulty stems from the fundamental disconnect between the compositional assumptions underlying most machine translation algorithms and the holistic, culturally-grounded meanings that characterize idiomatic language.

Idiomatic expressions in Arabic carry profound cultural meanings that frequently lack direct equivalents in English, making their translation an exceptionally complex linguistic endeavor that requires both semantic and cultural competence (Farghal & Haider, 2024). The Arabic language, shaped by centuries of diverse cultural influences including pre-Islamic Bedouin traditions, Islamic scholarship, and modern Arab nationalism, has developed a particularly sophisticated system of metaphorical expression. These idioms often draw from desert imagery, Islamic concepts, agricultural practices, and social customs that may be entirely foreign to English-speaking audiences, creating translation challenges that extend far beyond mere lexical substitution.

The challenge intensifies exponentially when machine translation systems attempt to process these culturally-embedded expressions, as they must navigate not only fundamental lexical differences but also intricate cultural and contextual nuances that often determine meaning (Haider & Shohaibar, 2024). Traditional neural machine translation systems, despite their sophisticated architectures, typically lack the cultural knowledge base necessary to recognize when literal translation approaches will fail and when alternative strategies become essential. This limitation becomes particularly pronounced when dealing with Arabic idioms that reference cultural practices, religious concepts, or historical events unfamiliar to English speakers.

The emergence of artificial intelligence-powered translation tools, particularly large language models like ChatGPT, has introduced revolutionary new possibilities for handling complex linguistic phenomena (Hidayati & Nihayah, 2024). These systems, trained on vast multilingual corpora and equipped with sophisticated attention mechanisms, demonstrate enhanced capabilities in contextual understanding and semantic processing. However, significant questions remain regarding their effectiveness compared to traditional neural machine translation systems like Google Translate, particularly in specialized domains requiring cultural sensitivity and contextual awareness. The fundamental question emerges: do these advanced AI systems represent genuine progress in cross-cultural communication, or do they merely offer more sophisticated ways of reproducing existing limitations?

This study addresses the critical gap in understanding how contemporary translation systems perform when confronting the specific challenge of translating Arabic idioms into English. While previous research has examined machine translation performance across various language pairs and linguistic phenomena, the particular challenges posed by Arabic idiomatic expressions in English translation contexts remain underexplored. This research gap is particularly significant given the increasing importance of Arabic-English communication in international business, diplomacy, academic exchange, and cultural dialogue.

Arabic idiomatic expressions function as sophisticated cultural repositories, encoding complex societal values, historical experiences, and collective wisdom within seemingly fixed linguistic structures (Shuhaiber & Haider, 2023). These expressions represent far more than linguistic curiosities; they constitute essential vehicles for cultural transmission and identity preservation. When Arabic speakers

employ idiomatic language, they invoke shared cultural understanding that extends beyond immediate semantic content to encompass historical memory, religious significance, and social relationships. This cultural depth creates translation challenges that purely linguistic approaches cannot adequately address.

When these culturally-rich expressions undergo machine translation, the preservation of both semantic content and cultural essence becomes paramount for achieving effective cross-cultural communication (Weld-Ali, Obeidat, & Haider, 2023). The translation process must somehow balance fidelity to source language cultural specificity with accessibility to target language audiences, a balancing act that requires sophisticated cultural and linguistic competence. Machine translation systems face the fundamental challenge of determining when cultural adaptation enhances communication and when it constitutes cultural erasure.

The theoretical foundations underlying idiom translation have been extensively developed in translation studies scholarship, with Baker's (1992) seminal work establishing key strategic categories that continue to influence contemporary research. However, the application of these theoretical frameworks to machine translation contexts reveals significant gaps between human translation decision-making processes and algorithmic approaches. While human translators can draw upon cultural knowledge, contextual awareness, and communicative purpose to guide strategic choices, machine translation systems must rely on statistical patterns and computational heuristics that may not capture essential cultural dimensions.

Furthermore, the evaluation of machine translation quality in idiomatic contexts presents unique methodological challenges that standard automatic evaluation metrics fail to capture adequately (Ali & Sayyiyed Al-Rushaidi, 2017). Traditional metrics focus on lexical similarity and structural correspondence, potentially missing crucial aspects of cultural appropriateness and communicative effectiveness that determine successful cross-cultural communication. This limitation necessitates more sophisticated evaluation approaches that can assess both linguistic accuracy and cultural preservation. The contemporary landscape of Arabic-English machine translation reflects broader tensions between technological advancement and cultural sensitivity (Abu-Ssaydeh, 2004). As machine translation systems become increasingly sophisticated and widely adopted, the stakes for accurate cross-cultural communication continue to rise. Inadequate translation of idiomatic expressions can lead to misunderstandings, cultural stereotyping, and communication failures with significant real-world consequences in diplomatic, business, and educational contexts.

Recent developments in large language model architectures have introduced new variables into this complex equation, as systems like ChatGPT demonstrate enhanced capabilities in contextual reasoning and cultural awareness (Khasawneh, 2023). However, these advances also raise new questions about consistency, reliability, and cultural appropriateness in automated translation systems. The fundamental challenge remains: can artificial intelligence systems develop genuine cultural competence, or are they limited to sophisticated pattern matching based on training data?

The significance of this research extends beyond purely academic concerns to address pressing practical needs in our increasingly interconnected global society. As Arabic-English communication becomes more frequent and consequential, understanding the capabilities and limitations of available translation technologies becomes essential for users ranging from international business professionals to academic researchers to diplomatic personnel. By providing detailed analysis of current system performance and strategic recommendations for improvement, this study aims to bridge the gap between technological capability and communicative need in cross-cultural contexts

This research investigates two fundamental questions:

1. How effectively do Google Translate and ChatGPT render Arabic idioms into English?
2. What translation strategies predominate in each system, and what implications do these strategies have for translation accuracy and cultural preservation?

Literature Review

Machine Translation Development

Machine translation has undergone a remarkable evolutionary journey since its inception, with neural machine translation representing the current state-of-the-art approach that has fundamentally transformed how automated systems process linguistic information (Zakraoui, Saleh, Al-Maadeed, & Alja'am, 2021). This technological progression reflects decades of computational linguistics research aimed at bridging the gap between human linguistic competence and machine processing capabilities. The journey from rule-based systems through statistical machine translation to contemporary neural architectures demonstrates both the persistent challenges and gradual advances in automated translation technology.

Google Translate, as one of the most widely deployed neural machine translation systems globally, has demonstrated significant improvements in handling complex linguistic structures, particularly in

morphologically rich languages like Arabic (Putera & Sujana, 2024). The system's neural architecture enables more sophisticated pattern recognition and contextual processing compared to earlier statistical approaches. However, substantial challenges persist with non-compositional language elements, where meaning cannot be derived from individual component analysis. This limitation becomes particularly pronounced when processing idiomatic expressions that require holistic semantic understanding rather than compositional interpretation.

The transition to neural machine translation architectures has brought both opportunities and limitations for Arabic-English translation contexts. While these systems demonstrate enhanced capability in capturing long-range dependencies and contextual relationships within sentences, they continue to struggle with culturally-embedded expressions that require knowledge beyond linguistic patterns (Zakraoui, Saleh, Al-Maadeed, & Alja'am, 2021). The architectural sophistication of neural networks, while impressive in computational terms, does not automatically translate to cultural competence or contextual sensitivity essential for effective idiom translation.

The integration of large language models like ChatGPT has introduced revolutionary conversational AI capabilities to translation tasks, offering unprecedented potential advantages in contextual understanding and cultural sensitivity (Li, 2024). These systems, trained on massive multilingual corpora using transformer architectures, demonstrate enhanced ability to recognize implicit meanings and cultural references that traditional neural machine translation systems might miss. The attention mechanisms underlying these models enable more sophisticated analysis of context and semantic relationships, potentially addressing some limitations of earlier translation approaches.

However, empirical evidence regarding large language model performance on specific linguistic phenomena, particularly idiomatic expressions, remains surprisingly limited despite widespread deployment and user adoption (Hidayati & Nihayah, 2024). This research gap creates uncertainty about the actual capabilities of these advanced systems in handling culturally-specific translation challenges. While anecdotal reports suggest improved performance, systematic evaluation of idiom translation quality across different cultural and linguistic contexts has not kept pace with technological development.

The architectural differences between traditional neural machine translation and large language models have important implications for idiom processing strategies. Traditional systems typically employ encoder-decoder architectures optimized for sentence-level translation, while large language models utilize more flexible attention mechanisms that can potentially capture broader contextual relationships essential for idiomatic interpretation (Li, 2024). These architectural distinctions may explain observed differences in translation approaches and quality, though comprehensive comparative analysis remains incomplete.

Arabic-English Translation Challenges

Arabic-English translation presents a constellation of unique challenges stemming from fundamental structural differences, profound cultural gaps, and complex directional asymmetries in idiomatic usage patterns (Ali & Sayyiyed Al-Rushaidi, 2017). These challenges extend far beyond simple lexical substitution to encompass deep-seated linguistic and cultural incompatibilities that complicate automated translation efforts. The historical and cultural distance between Arabic and English linguistic traditions creates translation obstacles that purely computational approaches struggle to address adequately.

Arabic idioms frequently employ sophisticated metaphorical constructions deeply rooted in desert culture, Islamic traditions, and Bedouin heritage, creating substantial translation complexities when these expressions must be rendered into English contexts that lack equivalent cultural frameworks (Abu-Ssaydeh, 2004). The metaphorical foundations of Arabic idiomatic language draw from experiential domains that may be entirely foreign to English-speaking audiences, including desert survival, camel herding, traditional Arabic poetry, and Islamic jurisprudence. These cultural references create translation challenges that extend beyond linguistic competence to require substantial cultural knowledge.

The structural characteristics of Arabic idiomatic expressions often reflect linguistic patterns that resist direct translation into English syntactic frameworks. Arabic's rich morphological system, flexible word order, and extensive use of metaphorical extension create idiomatic constructions that may appear semantically transparent to Arabic speakers while remaining opaque to machine translation systems trained primarily on surface-level pattern recognition (Ali & Sayyiyed Al-Rushaidi, 2017). The challenge intensifies when Arabic idioms employ linguistic devices like paronomasia, alliteration, and rhythmic patterns that carry meaning beyond semantic content.

Research has systematically identified several problematic areas in Arabic-English machine translation, including cultural reference preservation, metaphorical transfer accuracy, and semantic precision maintenance across cultural boundaries (Almahasees, 2021). Cultural reference preservation

presents particular difficulties when Arabic idioms invoke Islamic concepts, historical events, or social practices unfamiliar to English-speaking audiences. Machine translation systems must somehow balance cultural authenticity with communicative accessibility, a challenge that requires sophisticated cultural adaptation strategies.

These fundamental translation challenges become significantly amplified when dealing with idiomatic expressions that actively resist literal translation approaches and demand cultural mediation for effective communication (Almahasees, 2021). The non-compositional nature of idiomatic meaning creates obstacles for machine translation systems that rely primarily on statistical correlations and pattern matching rather than genuine semantic understanding. When Arabic idioms employ metaphorical frameworks incompatible with English conceptual systems, translation becomes a complex process of cultural negotiation rather than straightforward linguistic conversion.

The directional asymmetries between Arabic and English idiomatic systems further complicate translation efforts. While both languages possess rich idiomatic traditions, the cultural foundations underlying these traditions differ substantially, creating situations where direct equivalence becomes impossible and cultural adaptation becomes necessary (Abu-Ssaydeh, 2004). These asymmetries require translation strategies that go beyond linguistic competence to incorporate cultural mediation and pragmatic adaptation.

Translation Strategies for Idioms

Baker's (1992) seminal contribution established a foundational taxonomy for idiom translation strategies that continues to influence contemporary translation studies and provides essential analytical frameworks for evaluating machine translation approaches. Her systematic categorization includes: (1) translating with an equivalent idiom in the target language that maintains both semantic and formal correspondence, (2) translating with an idiom of similar meaning but different formal structure, (3) translation by paraphrase using non-idiomatic language to convey essential meaning, and (4) translation by omission when no adequate rendering proves possible. This comprehensive framework provides crucial analytical structure for evaluating machine translation performance and strategic consistency.

The application of Baker's framework to machine translation contexts reveals significant complexities that human translators navigate intuitively but that pose substantial challenges for automated systems (Salamah, 2015). Human translators can assess contextual appropriateness, audience expectations, and communicative purposes to guide strategic selections, while machine translation systems must rely on algorithmic heuristics that may not capture essential pragmatic considerations. This fundamental difference between human decision-making processes and machine processing approaches creates systematic challenges in strategy selection and implementation.

Subsequent research has expanded and refined these foundational categories to accommodate machine translation contexts, recognizing that automated systems may employ hybrid strategies or produce inconsistent results across semantically similar expressions (Salamah, 2015). The recognition of strategic inconsistency in machine translation outputs has important implications for user expectations and quality assessment. Unlike human translators who maintain strategic coherence across similar contexts, machine translation systems may apply different approaches to comparable expressions without systematic justification.

The application of translation strategies for Arabic idioms requires sophisticated cultural competence and contextual awareness that contemporary machine translation systems may fundamentally lack (Abu-Ssaydeh, 2004). Cultural competence involves understanding not only semantic relationships but also pragmatic appropriateness, cultural sensitivity, and audience expectations. Machine translation systems trained primarily on linguistic patterns may miss crucial cultural dimensions that determine strategic appropriateness and communicative effectiveness.

The challenge becomes particularly acute when Arabic idioms invoke religious concepts, historical references, or cultural practices that require substantial background knowledge for appropriate translation strategy selection. Baker's framework assumes translator competence in recognizing when equivalent idioms exist, when paraphrase becomes necessary, and when omission represents the most appropriate choice (Baker, 1992). Machine translation systems must somehow replicate these sophisticated judgment processes through computational means, a challenge that current technology has not fully resolved.

Contemporary research has identified additional strategic considerations specific to machine translation contexts, including the need for consistency maintenance across similar expressions and the importance of user expectation management (Salamah, 2015). These considerations extend beyond Baker's original framework to address unique challenges posed by automated translation deployment in diverse communicative contexts.

Empirical Studies in Machine Translation of Idioms

Recent empirical research has systematically examined machine translation performance on idiomatic expressions across various language pairs, revealing persistent patterns of strength and limitation that illuminate current technological capabilities and constraints. Motallebzadeh and Tousi (2011) conducted groundbreaking research demonstrating that sophisticated compensation strategies could substantially improve idiom translation quality while maintaining overall textual coherence and communicative effectiveness. Their findings suggested that strategic diversity in translation approaches could address some limitations of purely literal translation methods.

However, these same researchers identified fundamental obstacles preventing machine translation systems from implementing such sophisticated compensation approaches consistently across diverse idiomatic contexts (Motallebzadeh & Tousi, 2011). The inconsistency in strategic application represents a crucial limitation that distinguishes machine translation from human translation performance. While human translators can maintain strategic coherence across similar contexts, machine translation systems demonstrate problematic variability that undermines user confidence and communicative reliability.

Li and Chen (2019) conducted comprehensive comparative analysis of human and machine translation quality across multiple linguistic phenomena, reaching significant conclusions about current technological capabilities and limitations. Their research demonstrated that while AI translation systems have improved substantially in recent years, they have not yet achieved human-level performance in handling figurative language that requires cultural competence and contextual sensitivity. This performance gap represents a fundamental challenge rather than a temporary limitation, suggesting that current technological approaches may require substantial reconceptualization to achieve human-level cultural competence.

The gap between human and machine translation performance becomes particularly pronounced in culturally-specific expressions that require deep contextual understanding extending beyond surface-level linguistic analysis (Li & Chen, 2019). These findings have crucial implications for understanding the limitations of current machine translation technology and the continued necessity for human oversight in cross-cultural communication contexts. The persistence of this performance gap despite substantial technological advances suggests fundamental theoretical and methodological challenges that purely computational approaches may not resolve.

Khasawneh (2023) conducted important research investigating AI translation tools' role in facilitating cross-cultural communication, identifying persistent and systematic limitations in cultural nuance preservation that have significant implications for international communication effectiveness. This research revealed that while AI systems demonstrate enhanced linguistic processing capabilities, they continue to struggle with cultural adaptation requirements essential for effective cross-cultural communication. The findings suggest that technological sophistication in linguistic processing does not automatically translate to cultural competence.

The research identified specific areas where AI translation systems consistently fail to preserve cultural nuances essential for effective cross-cultural communication (Khasawneh, 2023). These limitations include inadequate recognition of cultural context, inappropriate cultural adaptation strategies, and inconsistent treatment of culturally-embedded expressions. Such findings suggest that current machine translation systems, despite impressive technological advances, require continued substantial development to handle idiomatic expressions with the cultural sensitivity necessary for effective international communication.

Recent research by Al-Salman and Haider (2024) has provided additional evidence of persistent challenges in AI-powered translation systems when handling culturally-specific content. Their comparative analysis of Google Translate, Gemini, and ChatGPT revealed significant variability in performance across different types of cultural content, with idiomatic expressions representing particularly challenging translation contexts. These findings reinforce the conclusion that cultural competence remains a fundamental limitation in current machine translation technology.

Contemporary Developments and Future Directions

The rapidly evolving landscape of machine translation technology has introduced new variables and possibilities that complicate traditional assessments of translation system capabilities. The integration of large language models into translation workflows represents a paradigm shift that may fundamentally alter how machine translation systems process culturally-embedded linguistic phenomena (Hidayati & Nihayah, 2024). However, empirical research has not yet provided comprehensive assessment of these new capabilities, creating uncertainty about their actual effectiveness in handling Arabic idiomatic expressions.

Recent studies have begun exploring the cultural sensitivity capabilities of advanced AI systems, with mixed results that suggest both promise and persistent limitations (Weld-Ali, Obeidat, & Haider,

2023). While these systems demonstrate enhanced contextual processing capabilities, fundamental questions remain about their ability to achieve genuine cultural competence rather than sophisticated pattern matching based on training data exposure. The distinction between cultural competence and cultural pattern recognition represents a crucial theoretical challenge with substantial practical implications for translation quality and cross-cultural communication effectiveness

Material and methods

This study employs a mixed-methods approach, combining quantitative frequency analysis with qualitative evaluation of translation accuracy and cultural preservation. The research design facilitates comprehensive assessment of machine translation performance across multiple dimensions.

Corpus Development

A corpus of 145 Arabic idiomatic expressions was compiled from diverse sources, including classical Arabic literature, contemporary media, and linguistic databases. Selection criteria prioritized expressions commonly used in modern Arabic discourse, ensuring relevance and representativeness. The corpus encompasses idioms from various semantic fields, including emotions, social relationships, time concepts, and cultural practices.

The following prompt was used with ChatGPT: "Translate the following Arabic idiomatic expressions and proverbs into English. Provide accurate and natural translations that preserve the meaning." Google Translate processed the same expressions through its standard interface.

Analytical Framework

The study adopts Baker's (1992) taxonomy as its foundational analytical framework, with modifications to accommodate machine translation contexts. Translation strategies are categorized as:

Literal Translation: Word-for-word rendering that may or may not preserve meaning

- (1A) Senseless literal translations
- (1B) Literal translations with lexical errors
- (1C) Literal translations with omissions
- (1D) Literal translations with other errors
- (1E) Meaningful literal translations

Sense-based Translation: Paraphrasing using common language

- (2A) Accurate sense transfer
- (2B) Partially accurate translations
- (2C) Inaccurate sense transfer

Figurative Language Translation: Using English idioms or metaphors

- (3A) Equivalent idiom translation
- (3B) Similar meaning, different form
- (3C) Non-idiomatic figurative language

Data Analysis

Quantitative analysis involves calculating frequencies and percentages of each translation strategy employed by both systems. Qualitative analysis examines selected examples to evaluate semantic accuracy, cultural preservation, and target language naturalness.

Results and discussion

Quantitative Analysis

The analysis reveals distinct patterns in translation strategy employment between Google Translate and ChatGPT. Table 1 presents comprehensive statistics for translation approaches used by both systems.

Table (1): Translation Strategy Statistics

Category	Sub-Category	Google Translate %	ChatGPT %
Literal Translation	Senseless word-for-word	45%	28%
	Lexical errors	14%	3%
	Omissions	3%	4%
	Other errors	2%	1%
	Meaningful literal	10%	12%
	Total Literal	74%	48%
Sense-based Translation	Accurate transfer	12%	31%
	Partially accurate	2%	6%
	Inaccurate transfer	1%	4%
	Total Sense-based	15%	41%
Figurative Translation	Equivalent idiom	8%	9%
	Similar meaning	2%	1%
	Non-idiomatic figurative	1%	1%

Category	Total Figurative	11%	11%
	Sub-Category	Google Translate %	ChatGPT %

When comparing the translation tendencies of Google Translate and ChatGPT, a clear difference in approaches becomes evident. Google Translate demonstrates a strong bias toward literal translation strategies, with approximately 74% of its renditions falling into this category. While such an approach may sometimes preserve the structural elements of the source text, it often comes at the expense of naturalness and clarity. In fact, the majority of these literal translations, about 45%, are nothing more than mechanical, word-for-word renderings that frequently fail to capture the intended sense of the original passage. As a result, the target text can appear awkward, fragmented, or even nonsensical to the reader, reducing its communicative effectiveness.

By contrast, ChatGPT displays a more balanced translation profile. Although it still employs literal translation in 48% of cases, it also shows a considerably higher reliance on sense-based strategies, accounting for 41% of its output. This indicates that ChatGPT is more flexible in interpreting context and meaning, striving not only to replicate the linguistic form but also to convey the semantic content more naturally. The greater variation in approach allows ChatGPT to generate translations that are often more comprehensible, contextually accurate, and communicatively appropriate than those produced by Google Translate.

Qualitative Analysis

Literal Translation Challenges

Both systems exhibit persistent issues with literal translation approaches. Consider these examples:

Table (2): Literal Translation Examples

Arabic Idiom	Meaning	Google Translate	ChatGPT	Translation Type
كلامه عسل	His words are sweet/persuasive	His words honey	His speech is honey	Senseless literal
ضرب عصفورين بحجر واحد	Kill two birds with one stone	Hit two birds with one stone	Strike two birds with one stone	Meaningful literal
الصبر مفتاح الفرج	Patience is the key to relief	Patience key relief	Patience is the key to relief	Literal with omission

The first example demonstrates how literal translation fails to capture the metaphorical meaning of persuasive speech. Both systems translate the components without recognizing the idiomatic nature. The second example shows successful literal translation where the Arabic idiom closely parallels an English equivalent. The third reveals omission errors that compromise grammatical accuracy.

- **Sense-based Translation Success and Failures**

ChatGPT demonstrates superior performance in sense-based translations, though inconsistencies remain:

Table (3): Sense-based Translation Examples

Arabic Idiom	Meaning	Google Translate	ChatGPT	Accuracy
نام على حرير	He lived luxuriously	Slept on silk	He lived in luxury	Accurate (ChatGPT)
أكل الهوا	He was disappointed	Ate air	He was let down	Accurate (ChatGPT)
شال الهم	He was relieved	Lifted worry	He felt relieved	Accurate (ChatGPT)
ركب الموجة	He followed the trend	Rode the wave	He jumped on the bandwagon	Accurate (ChatGPT)

These examples illustrate ChatGPT's capability to recognize idiomatic meanings and provide appropriate English equivalents. Google Translate consistently applies literal strategies, missing the figurative intentions.

- **Cultural Context Preservation**

A critical challenge emerges in preserving cultural contexts while achieving target language naturalness:

Table (4): Cultural Context Examples

Arabic Idiom	Cultural Context	Google Translate	ChatGPT	Cultural Preservation
حج ولا زاره	Religious pilgrimage metaphor	Pilgrimage not visit	All show, no substance	Lost (ChatGPT better)
باع الهوى في قرب	Desert trade metaphor	Sold air in bags	He sold false promises	Moderate (ChatGPT)
ماء وطين	Basic elements metaphor	Water and mud	Complete chaos	Lost (ChatGPT better)

While ChatGPT provides more natural English expressions, cultural specificity often disappears in the translation process. This represents a fundamental challenge in cross-cultural idiom translation.

Discussion

Quantitative Findings and Their Implications

The quantitative results from this study provide compelling confirmation of previous research findings regarding persistent machine translation limitations when confronting idiomatic expressions, particularly in the complex Arabic-English translation context (Almahasees, 2021). These findings reveal systematic patterns that extend beyond individual system performance to illuminate fundamental challenges in current machine translation approaches to culturally-embedded linguistic phenomena. The statistical patterns observed in this research demonstrate remarkable consistency with established theoretical predictions while also revealing unexpected aspects of contemporary translation system behavior.

Google Translate's pronounced reliance on literal translation approaches, evidenced by the 74% literal translation rate observed in this study, aligns remarkably well with established neural machine translation tendencies previously identified by researchers examining similar linguistic phenomena across multiple language pairs. This consistency suggests that the challenges observed in Arabic-English idiom translation represent broader limitations in neural machine translation architectures rather than language-specific anomalies. The systematic preference for compositional analysis over holistic semantic interpretation reflects fundamental architectural constraints that persist despite technological advances.

However, the extent of literal translation usage documented in this study (74%) substantially exceeds reasonable expectations based on previous research and suggests more severe limitations in idiom recognition capabilities than previously documented (Almahasees, 2021). This finding indicates that Google Translate's neural architecture, despite sophisticated attention mechanisms and extensive training data, continues to struggle with fundamental idiom identification processes essential for appropriate translation strategy selection. The high frequency of literal translation attempts suggests systematic failure in the preliminary step of recognizing when expressions require non-compositional interpretation approaches.

The implications of this finding extend beyond technical performance metrics to encompass serious concerns about user experience and communication effectiveness. When machine translation systems fail to recognize idiomatic expressions and apply inappropriate literal translation strategies, the resulting outputs often confuse rather than clarify intended meanings, potentially creating cross-cultural misunderstandings with significant real-world consequences (Abu-Ssaydeh, 2004). This systematic tendency toward inappropriate literal translation represents a fundamental reliability concern that users must consider when deploying these systems for cross-cultural communication purposes.

ChatGPT's more balanced approach to translation strategy selection reflects the significant potential advantages of large language models in contextual understanding and semantic processing capabilities (Hidayati & Nihayah, 2024). The system's 41% sense-based translation rate represents a substantial improvement compared to traditional neural machine translation approaches, indicating enhanced semantic processing capabilities that may address some fundamental limitations of earlier translation architectures. This improvement suggests that large language model architectures may offer promising directions for addressing persistent challenges in idiomatic expression translation.

The enhanced performance observed in ChatGPT's approach appears to stem from the system's more sophisticated attention mechanisms and broader contextual processing capabilities, which enable recognition of semantic relationships that traditional neural machine translation systems might miss (Li, 2024). The ability to process longer contextual windows and maintain attention across extended sequences may contribute to improved idiom recognition and more appropriate translation strategy

selection. These architectural advantages represent genuine progress in addressing fundamental challenges that have persistently limited machine translation effectiveness.

However, the persistence of literal translations at a substantial rate (48%) in ChatGPT outputs suggests that even advanced large language model architectures face continued challenges in idiom identification and cultural competence development (Hidayati & Nihayah, 2024). This finding indicates that technological sophistication alone may not resolve fundamental challenges related to cultural knowledge representation and contextual sensitivity. The substantial literal translation rate suggests that current large language model training approaches may not adequately address cultural competence requirements essential for effective idiomatic expression translation.

Qualitative Analysis and Cultural Context Preservation

The qualitative analysis conducted in this study reveals fundamental and persistent issues in cultural context preservation that extend far beyond surface-level linguistic accuracy to encompass deeper challenges in cross-cultural knowledge representation and cultural sensitivity (Weld-Ali, Obeidat, & Haider, 2023). These findings support and extend previous observations regarding the particular difficulties that machine translation systems encounter when processing religious and cultural expressions that carry meanings extending beyond immediate semantic content to encompass cultural identity, historical memory, and social relationships.

Both translation systems examined in this study demonstrate systematic struggles to maintain cultural specificity while simultaneously achieving target language naturalness, reflecting broader theoretical and practical challenges inherent in cross-cultural communication mediated by computational systems (Khasawneh, 2023). This fundamental tension between cultural authenticity and communicative accessibility represents one of the most persistent challenges in machine translation development, requiring sophisticated cultural adaptation strategies that current systems have not successfully implemented.

The qualitative evidence reveals that machine translation systems often prioritize linguistic fluency over cultural accuracy when these two goals conflict, resulting in translations that may read naturally in English while completely erasing essential cultural content from the source expressions (Shuhaiber & Haider, 2023). This systematic bias toward target language naturalization reflects training data patterns and optimization objectives that may inadvertently promote cultural homogenization rather than cultural preservation. The implications of this bias extend beyond technical performance to encompass important questions about cultural representation and preservation in global communication contexts.

Furthermore, the analysis reveals systematic inconsistencies in cultural adaptation approaches, with both systems applying different strategies to culturally similar expressions without apparent logical justification (Salamah, 2015). This inconsistency undermines user confidence and creates unpredictable results that complicate effective cross-cultural communication planning. The lack of systematic cultural adaptation principles suggests that current machine translation approaches may require fundamental reconceptualization to address cultural competence requirements effectively.

The specific challenges observed in translating idioms rooted in Islamic traditions highlight the particular difficulties that machine translation systems face when processing religious and spiritual content that requires deep cultural knowledge for appropriate interpretation (Weld-Ali, Obeidat, & Haider, 2023). These expressions often carry multiple layers of meaning that extend beyond immediate semantic content to encompass theological concepts, historical references, and spiritual practices that may be entirely unfamiliar to target language audiences. The systematic failure to preserve these cultural dimensions represents a significant limitation in current machine translation capabilities.

Technological Advances and Persistent Limitations

The superior performance of ChatGPT in sense-based translations, as documented in this study, aligns closely with recent research findings regarding large language model advantages in semantic processing and contextual understanding (Li, 2024). These improvements represent genuine technological progress that addresses some fundamental limitations of earlier neural machine translation approaches. The enhanced ability to recognize implicit meanings and process extended contextual relationships demonstrates the potential of advanced AI architectures to overcome traditional machine translation limitations.

However, the inconsistency in translation quality observed across semantically similar expressions indicates that these apparent improvements may not represent systematic advancement in cultural competence but rather statistical variations in training data exposure and processing approaches (Li & Chen, 2019). This finding raises important questions about the reliability and consistency of large language model performance in specialized translation contexts. The variability in performance across similar expressions suggests that current training approaches may not adequately ensure consistent cultural competence development.

The technological advances represented by large language models bring both opportunities and new challenges for cross-cultural communication applications. While these systems demonstrate enhanced contextual processing capabilities, they also introduce new forms of unpredictability and inconsistency that complicate user expectations and deployment strategies (Putera & Sujana, 2024). The enhanced capabilities come with increased complexity that may require new approaches to quality assessment and user guidance.

The comparison between Google Translate and ChatGPT performance reveals important insights about the relationship between architectural sophistication and cultural competence in machine translation systems. While ChatGPT's advanced architecture provides advantages in semantic processing, the persistent challenges in cultural context preservation suggest that architectural improvements alone may not resolve fundamental cultural competence limitations (Khasawneh, 2023). These findings indicate that cultural competence development may require specialized training approaches and knowledge representation strategies that current general-purpose language model training does not provide.

Cultural Preservation as a Critical Concern

Cultural preservation emerges from this analysis as a critical concern that extends beyond technical performance metrics to encompass fundamental questions about cultural representation and identity preservation in global communication contexts. The systematic challenges observed in translating idioms rooted in Islamic traditions or Arab heritage highlight the particular vulnerabilities of minority cultural perspectives in machine translation systems trained primarily on dominant cultural patterns (Haider, Saideen, & Hussein, 2023).

The tendency observed in both translation systems to prioritize target language naturalness over cultural authenticity raises profound questions about translation purposes, audience expectations, and the appropriate balance between accessibility and authenticity in cross-cultural communication (Al-Salman & Haider, 2024). This systematic bias toward cultural adaptation reflects optimization objectives and training data patterns that may inadvertently promote cultural homogenization rather than cultural diversity preservation. The implications extend beyond technical considerations to encompass important ethical questions about cultural representation in global communication technologies.

The systematic erasure of cultural specificity observed in many translation outputs represents a form of cultural loss that may have significant implications for cultural identity preservation and cross-cultural understanding development (Almrayat, Farghal, & Haider, 2024). When machine translation systems consistently remove cultural markers in favor of generic expressions, they may contribute to cultural flattening that reduces the richness and diversity of cross-cultural communication. This cultural flattening may impede rather than facilitate genuine cross-cultural understanding development.

The analysis reveals particular challenges in preserving religious and spiritual content that carries deep cultural significance for Arabic-speaking communities. The systematic failure to maintain these cultural dimensions suggests that current machine translation approaches may require fundamental reconceptualization to address cultural preservation requirements effectively (Farghal & Haider, 2024). The development of culturally-sensitive translation approaches may require interdisciplinary collaboration between computational linguists, cultural experts, and community representatives to ensure appropriate cultural representation.

Implications for Cross-Cultural Communication

The findings of this study have significant implications for cross-cultural communication practices and the deployment of machine translation technologies in international contexts. The persistent limitations in cultural context preservation documented in this research suggest that current machine translation systems may not be adequate for cross-cultural communication contexts where cultural sensitivity and authenticity are paramount (Weld-Ali, Obeidat, & Haider, 2023).

The systematic challenges in idiom translation revealed by this analysis indicate that users must exercise considerable caution when deploying machine translation systems for cross-cultural communication purposes, particularly in contexts involving cultural or religious content (Khasawneh, 2023). The unpredictable quality variations observed across similar expressions suggest that machine translation outputs require careful human review and cultural validation to ensure appropriate cross-cultural communication effectiveness.

The research findings also highlight the continued necessity for human cultural expertise in translation processes, particularly when dealing with culturally-embedded expressions that require deep contextual knowledge for appropriate interpretation (Li & Chen, 2019). The technological advances represented by large language models, while impressive, have not eliminated the need for human cultural competence in translation quality assurance and cultural adaptation processes.

Theoretical Implications and Future Directions

The theoretical implications of this research extend beyond immediate technical performance concerns to encompass fundamental questions about the nature of cultural competence in artificial intelligence systems and the appropriate role of machine translation in cross-cultural communication contexts (Baker, 1992). The persistent challenges in cultural context preservation suggest that current computational approaches to translation may require substantial theoretical reconceptualization to address cultural competence requirements effectively.

The findings indicate that future machine translation development should prioritize cultural competence development alongside linguistic accuracy improvement, requiring interdisciplinary approaches that combine computational linguistics expertise with cultural knowledge and community input (Motallebzadeh & Tousi, 2011). The development of culturally-sensitive translation approaches may require new theoretical frameworks that explicitly address cultural representation and preservation requirements.

The research also suggests that evaluation metrics for machine translation quality should incorporate cultural preservation assessments alongside traditional linguistic accuracy measures, requiring new methodological approaches that can assess cultural appropriateness and authenticity (Ali & Sayyiyed Al-Rushaidi, 2017). The development of culturally-sensitive evaluation approaches represents an important direction for future research in machine translation assessment and improvement.

Conclusion

This study reveals significant disparities in Arabic-to-English idiom translation performance between Google Translate and ChatGPT. While ChatGPT demonstrates advantages in semantic processing and sense-based translation, both systems face fundamental challenges in cultural context preservation and consistent idiom recognition.

The persistence of literal translation approaches, particularly in Google Translate, confirms that current machine translation technology has not fully mastered non-compositional language elements. The high frequency of senseless literal translations (45% in Google Translate, 28% in ChatGPT) indicates systematic failures in idiom identification processes.

The study's findings have implications for machine translation development and cross-cultural communication practices. The superior performance of ChatGPT in sense-based translations suggests that large language models offer promising directions for improving idiom translation quality. However, the lack of consistency and cultural preservation issues indicate that current systems require substantial development before achieving human-level performance.

Future research should examine the role of cultural competence in machine translation systems and investigate methods for preserving cultural specificity while maintaining target language naturalness. Additionally, longitudinal studies tracking machine translation improvement across specific linguistic phenomena would provide valuable insights into technological development trajectories.

The results indicate that while machine translation has advanced considerably, human oversight remains essential for accurate cross-cultural communication, particularly when dealing with culturally-specific idiomatic expressions. The gap between technological capability and cultural competence represents a fundamental challenge requiring interdisciplinary approaches combining linguistic expertise with artificial intelligence development.

Machine translation systems must evolve beyond lexical and syntactic processing to incorporate cultural knowledge and contextual sensitivity. Until such advances occur, users should approach machine-translated idiomatic content with caution, particularly in contexts where cultural accuracy is paramount.

References

[1] Abu-Ssaydeh, A.-F. (2004). Translation of English idioms into Arabic. *Babel. Revue internationale de la traduction / International Journal of Translation*, 50(2), 114-131.

[2] Ali, H., & Sayyiyed Al-Rushaidi, S. M. (2017). Translating idiomatic expressions from English into Arabic: Difficulties and strategies. *Arab World English Journal*, 1(3), 22-31.

[3] Almahasees, Z. (2021). *Analysing English-Arabic Machine Translation: Google Translate, Microsoft Translator and Sakhr*. London: Routledge.

[4] Almrayat, O., Farghal, M., & Haider, A. S. (2024). Translating Idioms in Political Discourse: Translation Strategies and Semantic Categories. *International Journal of Arabic-English Studies*, 25(1), 187-210.

[5] 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), 2483-2498.

[6] Baker, M. (1992). *A coursebook on translation*. London and New York: Routledge.

[7] Farghal, M., & Haider, A. S. (2024). Translating classical Arabic verse: human translation vs. AI large language models (Gemini and ChatGPT). *Cogent Social Sciences*, 10(1), 2410998.
[8] Haider, A. S., & Shohaibar, R. S. (2024). Netflix English subtitling of idioms in Egyptian movies: challenges and strategies. *Humanities & Social Sciences Reviews*, 11(1), 949.
[9] Haider, A. S., Saideen, B., & Hussein, R. F. (2023). Subtitling Taboo Expressions from a Conservative to a More Liberal Culture: The Case of the Arab TV Series Jinn. *Middle East Journal of Culture and Communication*, 16(4), 363-385.
[10] Hidayati, N. N., & Nihayah, D. H. (2024). Google Translate, ChatGPT or Google Bard AI: A Study toward Non-English Department College Students' Preference and Translation Comparison. *Inspiring: English Education Journal*, 7(1), 14-33.
[11] Khasawneh, M. A. S. (2023). The Potential of AI in Facilitating Cross-Cultural Communication Through Translation. *Journal of Namibian Studies: History Politics Culture*, 37, 107-130.
[12] Li, H., & Chen, H. (2019). Human vs. AI: An assessment of the translation quality between translators and machine translation. *International Journal of Translation, Interpretation, and Applied Linguistics (IJTIAL)*, 1(1), 1-12.
[13] Li, X. (2024). Comparison of Translation Quality between Large Language Models and Neural Machine Translation Systems: A Case Study of Chinese-English Language Pair. *International Journal of Education and Humanities*, 4(2), 121-128.
[14] Motallebzadeh, K., & Tousi, S. (2011). Employing compensation strategy in Translation of Idioms: A case study of the Translation of Mark Twain's Adventures of Huckleberry Finn in Persian. *International Journal of Linguistics*, 3(1), 1-9.
[15] Putera, L. J., & Sujana, I. M. (2024). A comparative analysis of accuracy between Google Translate and Bard in translating abstracts of scientific journals. *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, 10(1), 35-44.
[16] Salamah, D. A. (2015). An investigation of idiom comprehension and translation by translation students at KSU. *Arab World English Journal*, 6(3), 293-308.
[17] Shuhaiber, R., & Haider, A. S. (2023). Strategies of Subtitling Egyptian Idiomatic Expressions and Proverbs into English. *Language Value*, 16(2), 60-99.
[18] Weld-Ali, E. W., Obeidat, M. M., & Haider, A. S. (2023). Religious and Cultural Expressions in Legal Discourse: Evidence from Interpreting Canadian Courts Hearings from Arabic into English. *International Journal for the Semiotics of Law-Revue internationale de Sémiotique juridique*, 36(6), 2283–2301.
[19] Zakraoui, J., Saleh, M., Al-Maadeed, S., & Alja'am, J. M. (2021). Arabic Machine Translation: A Survey With Challenges and Future Directions. *IEEE Access*, 9, 161445-161468.