

Beyond Lexical Equivalence: A Context-Sensitive Translation Model for Evaluating ChatGPT English Translation of Culturally Loaded Arabic Expressions

MANAL MAHMOUD ALZARIENI¹

Abstract

Recent advances in artificial intelligence (AI) have substantially improved machine translation. However, the translation of culturally loaded expressions remains challenging in that such expressions require semantic, pragmatic and cultural interpretation beyond lexical equivalence. Existing research has primarily evaluated AI translation quality by comparing AI outputs with human translations across specific genres. Nevertheless, limited attention has been given to developing practical frameworks for assessing and improving AI translation of culturally embedded Arabic expressions. This study investigates the ability of ChatGPT to translate culturally loaded Arabic expressions into English and proposes a Context-Sensitive Translation Model (CSTM) for evaluating and improving AI-assisted translation. Grounded in translation equivalence theory and principles of culture-specific translation, the study adopts a qualitative descriptive design. A purposive sample of 16 Arabic cultural terms which represent eight cultural categories was translated using ChatGPT and then analyzed through three evaluative dimensions, namely, semantic, pragmatic and cultural equivalence. The findings obtained reveal that ChatGPT consistently produces fluent and semantically acceptable translations. However, it frequently fails to preserve the deeper cultural, legal, religious and social meanings embedded in the source expressions. Although ChatGPT successfully conveyed the general lexical meaning of the selected terms, it could only partially maintain their pragmatic and cultural significance. The study demonstrated that lexical accuracy alone is insufficient for culturally appropriate translation. The study proposes CSTM as a practical framework for evaluating AI-generated translations and recommends integrating semantic, pragmatic and cultural considerations into AI translation systems while maintaining human involvement in translating culturally sensitive texts.

Keywords: *AI Translation, Arabic-English Translation, Context-Sensitive Translation Model, Culturally Loaded Expressions, Translation Equivalence.*

Introduction

With the help of recent advances in AI, translation practice has become easier and more efficient; more powerful neural machine translation (NMT) systems and large language models (LLMs) can instantly produce fluent multilingual texts. These systems have been increasingly employed in cross-cultural communication as well as professional translation; examples of these systems include DeepL, ChatGPT, Google Translate and Microsoft Copilot. Such AI systems are able to process large volumes of linguistic data and produce contextually appropriate translation which all enhanced the efficacy as well as accessibility of translation. Accordingly, several AI users usually depend on these systems for both daily communication and domain-specific translations (Koehn, 2020).

However, there has been doubt regarding the extent to which such AI systems are capable of translating texts which include culturally loaded terms. Literature (e.g. Azizov, 2024) reveals that current AI models can recognize most phrases and produce literal translations but find it difficult to reflect cultural meanings of cultural expressions from one language to another. This is because such expressions reflect shared cultural knowledge, religious beliefs, historical realities and social practices and values that cannot be fully understood without reference to their specific sociocultural context (Newmark, 1988; Baker, 2018). Therefore, statistical or neural pattern-based translations can provide grammatically correct translations but often fail to convey the holistic cultural meanings of a given text.

¹ Assistant Professor, Department of English Language and Literature, Faculty of Arts and Sciences, The World Islamic Sciences & Education University (WISE), Email: manal.alzarieni@wise.edu.jo, ((Corresponding author))

When it comes to translating culturally and linguistically complex languages, Arabic is definitely considered as one of the most difficult languages for AI. This is because the Arabic has a large inventory of culturally specific expressions that are related to tribal customs, Islamic beliefs and traditions, kinship relationships, hospitality, social etiquette and history. Terms like *صدقة sadaqah* (voluntary charity) are culturally bound to Arab and Islamic culture; such cultural terms do not have equivalents in English. Their interpretation depends on both dictionaries as well as context and pragmatic reasoning about how language and culture are understood by interlocutors in the source community. Literal translations of these expressions almost certainly result in semantic loss, cultural misrepresentation or complete obliteration of pragmatic content, thus restricting intercultural communication (Al-Qinai, 2012; Dickins, Hervey & Higgins. 2017).

Researchers have long wondered how to deal with the issues of translating cultural terms. In this regard, Newmark (1988) claims that translation is not replacing one language with another but negotiating meaning from one culture to another. Moreover, Aixelá (1996) and Pedersen (2011) claimed that the accurate translation of culture-specific items involves the use of appropriate translation strategies like borrowing, cultural substitution, adaptation, explanation, or descriptive translation based on the communicative purpose and target readership. Recent studies on NMT show that even though the overall translation quality of AI systems is better than that of earlier statistical approaches. AI systems still encounter challenges while translating idioms, religious terminology, cultural terms and pragmatic expressions among other items that demand reasoning beyond lexical correspondence (Castilho et al., 2017; Freitag et al., 2021).

The emergence of large language models has renewed optimism regarding AI-assisted translation. Unlike conventional NMT systems, LLMs are capable of generating translations by integrating broader contextual information and demonstrating stronger discourse-level coherence. Recent research suggests that such models can produce more natural translations and respond to user prompts that specify audience, style or communicative purpose (Jiao et al., 2023; Hendy et al., 2023). However, evidence also indicates that their performance remains inconsistent when translating culturally loaded expressions. Although AI systems can produce fluent English equivalents, they do not sometimes reflect implicit cultural meanings in translation and often overgeneralize concepts or introduce culturally inappropriate interpretations that change the communicative intention of the original text (Al-Jarf, 2025).

Current studies have predominantly paid attention to evaluating the linguistic accuracy of AI-generated translations using automatic evaluation metrics like COMET, BLEU and METEOR or through comparisons with human translations. While research provides valuable insights into overall translation quality, it often pays less attention to cultural equivalence as well pragmatic meaning. In particular, few or no studies have systematically investigated how AI systems translate culturally loaded Arabic terms into English or proposed practical frameworks that can guide AI towards culturally appropriate translation choices. Current research tends to identify translation errors without offering principled solutions for enhancing AI-assisted translation in culturally sensitive contexts.

This gap is vitally significant in that culturally loaded expressions frequently occur in legal, religious, literary, media and governmental texts where inaccurate translation may lead to misunderstanding as well as cultural misrepresentation. Because AI-generated translations have increasingly been integrated into professional translation practice, improving the treatment of culturally specific language has become a significant challenge for both translation studies as well as AI research. Recent evidence, rather than treating AI as a substitute for human translators, suggests that the solution lies in a more human-centred approach to AI where AI supports translators while cultural interpretation and contextual decision-making remain guided by human expertise (Bowker, 2023).

Against this background, the current study examines how recent AI systems translate culturally loaded Arabic terms into English. It investigates whether and how the extent to AI translations retains semantic, pragmatic and cultural meanings across general categories of culture-specific expressions. Based on the previously identified translation patterns and recurring limitations, the present study proposes CSTM designed to improve the quality of AI-assisted translation of culturally embedded Arabic expressions. By integrating principles from translation studies with recent developments in AI-assisted language

technologies, the proposed model in the current study aims at bridging the gap between computational translation efficiency as well as culturally informed human interpretation. Consequently, this study contributes to current debates on culturally aware AI, Arabic-English translation, and the evolving relationship between AI systems and human translators in multilingual communication.

Literature Review

AI translation systems have recently become common and thus interest has increased in evaluating the AI translation systems deal with culture-bound language. Although LLMs and NMT have enhanced translation accuracy and linguistic fluency, studies (e.g. Vishwakarma, 2023) have revealed that culturally loaded terms pose challenges for translators in that such terms require pragmatic, semantic and cultural interpretation alongside lexical equivalence. As a result, the focus of many of these studies (e.g. Al-Khresheh, 2025) has shifted towards investigating the extent to which AI can precisely preserve culturally embedded meanings in Arabic-English translation.

Much research has compared AI-generated translations of culturally rich Arabic texts with those of human translators. For instance, AlAfnan and Alshakhs (2025) revealed that AI systems generated fluent translations of Arabic dialect poetry. However, they failed to preserve poetic rhythm, metaphorical meaning, and cultural symbolism while human translators showed greater interpretive sensitivity. Similarly, Hanan and Sukayna (2025) found that ChatGPT translated expressions from Naguib Mahfouz's *The Cairo Trilogy* coherently but encountered challenges while translating religious references, metaphors and culturally nuanced expressions. Similarly, Mehassouel and Djeha (2024) pointed out that Gemini outperformed ChatGPT in translating culture-specific items. Nevertheless, they concluded that both systems still required human post-editing to guarantee cultural accuracy. Such studies consistently recommend that there should be a collaborative human-AI translation model where AI functions as a supportive rather than a replacement technology.

Other studies have analyzed AI performance across various types of culturally loaded expressions. For example, Ennouari and Houssaini (2025) demonstrated that machine translation usually failed to retain the authenticity of Arabic proverbs and anecdotes. Al-Jarf (2025) also demonstrated that both Microsoft Copilot and student translators heavily relied on literal translation while translating expressions of impossibility. Similarly, AlGhamdi and Ahmed Al-Sofi (2025) observed that Gemini translated culture-specific terms on Saudi tourist signs more accurately compared to DeepSeek though human revision remained essential. Alkhatnai (2025) and Al Maaytah (2026) presented similar conclusions and reported that AI systems perform effectively on literal texts but still experience difficulties with idiomatic language, context-dependent meanings and cultural nuances.

Finally, previous studies illustrate that AI translation has reached a high level of linguistic sophistication but remains constrained in interpreting culturally embedded meanings. Most current research, however, focuses on assessing translation quality within specific genres like poetry, literature, idioms or tourist signs, and mainly compares AI-generated translations with human translations. Less attention has been given to developing practical frameworks that enhance AI translation of culturally loaded Arabic expressions. Therefore, the current study fills this research gap by examining a broader range of culturally embedded Arabic terms and proposing a context-sensitive translation model that incorporates semantic, pragmatic and cultural considerations to improve AI-assisted Arabic-English translation.

Methodology

Study Design

This study adopts a qualitative descriptive research design to examine how contemporary AI translation systems translate Arabic culturally loaded expressions into English. A qualitative approach is appropriate as the study aims to investigate how semantic, pragmatic and cultural meanings are preserved or lost in translation. Different from studies which primarily depend on automatic evaluation measures like COMET

or BLEU, the present study focuses on analyzing the quality of AI-generated translations within their cultural and communicative contexts. Moreover, the study also adopts a comparative analytical approach by investigating the translations of ChatGPT to identify recurring translation patterns, common errors and differences in their treatment of culturally embedded expressions. The findings subsequently inform the development of a context-sensitive translation model designed to enhance AI-assisted translation of culturally loaded Arabic expressions.

Data Collection and Sampling

The data of the present study include 16 culturally loaded Arabic expressions that were purposively selected; the sample selected represents various aspects of Arab and Islamic culture. The sample consists of expressions related to religion, social norms, hospitality, legal concepts, governance and institutions, family values, tribal culture and social relationships. Expressions were identified from authoritative Arabic dictionaries, cultural reference works and authentic Arabic texts to ensure their cultural relevance as well as frequent use.

This sample is selected because the study objective is not statistical representativeness but the in-depth investigation of culturally significant expressions which pose challenges for AI translation. The selection of data was guided by three criteria. First, each expression contains a cultural meaning that extends beyond its literal lexical interpretation. Second, the expression lacks a direct one-to-one lexical equivalent in English and thus requires cultural interpretation during translation. Third, the selected expressions represent a number of cultural categories to guarantee comprehensive coverage of different translation challenges. Each Arabic expression was independently translated using selected ChatGPT 5.2. The resulting English translations constituted the primary dataset for qualitative analysis.

Table 1. Selected Cultural Terms

Cultural categories	Selected cultural terms
Religious concepts	بركة <i>barakah</i> (blessing), وقف <i>waqf</i> (endowment)
Islamic legal concepts	عدة <i>iddah</i> (waiting period), محرم <i>mahram</i> (guardian)
Governance and institutions	شورى <i>shura</i> (consultation), ديوان <i>diwan</i> (council)
Hospitality	مجلس <i>majlis</i> (sitting room), ضيافة <i>diyafah</i> (hospitality)
Family values	صلة الرحم <i>silat al-Rahim</i> (maintaining family ties); قوامه <i>qiwamah</i> (guardianship)
Social norms	العيب <i>al-'Ayb</i> (shame); فزعة <i>faẓ'ah</i> (helping others)
Tribal culture	نخوة <i>nakhwa</i> (chivakry), ثأر <i>tha'r</i> (revenge)
Social relationships	وساطة <i>wasta</i> (connections); تكافل <i>takāful</i> (solidarity)

Theoretical Foundations of the Context-Sensitive Translation Model

This subsection explains that CSTM is not arbitrarily created. Instead, it incorporates well-established concepts from translation studies into a single framework specifically designed for the assessment of AI-generated translations of culturally loaded Arabic expressions.

TABLE. 2 Theoretical foundations of CSTM

CSTM Dimension	Theoretical Basis
Semantic Equivalence	Nida (1964); Baker (2018)
Pragmatic Equivalence	Baker (2018); House (2015)
Cultural Equivalence	Newmark (1988); Aixelá (1996)

CSTM is theoretically grounded in established approaches to translation equivalence rather than representing an entirely new theory of translation. It synthesizes three complementary dimensions, namely, semantic, pragmatic and cultural equivalence which are derived from influential work in translation studies.

Semantic equivalence primarily draws on the notion of equivalence proposed by Nida (1964) as well as Baker's (2018) discussion of meaning across languages. In addition, pragmatic equivalence is informed by treatment of implicature (Baker, 2018) and communicative meaning together with the functional approach to translation quality assessment proposed by House (2015). Cultural equivalence builds on Newmark's (2018) treatment of cultural words and Aixelá's (1996) work on culture-specific items. By integrating those perspectives into a single analytical framework, the proposed CSTM offers a systematic method for the evaluation of AI-generated translations of culturally loaded Arabic expressions.

Method of Data Analysis

The selected expressions were all analyzed through the proposed model i.e. Context-Sensitive Translation Model (CSTM). This model evaluates the quality of AI-generated translations according to three complementary dimensions: semantic equivalence, pragmatic equivalence and cultural equivalence. All translated expressions were assessed against the original Arabic expressions to determine how many of those dimensions were preserved.

In the present study, semantic equivalence refers to the extent to which denotative meaning of a cultural term is transferred through its equivalent. In addition, pragmatic equivalence examines whether a translation maintains the communicative function and contextual use of its original expression. Cultural equivalence is concerned with the extent to which the translation succeeds in providing equivalent and comparable cultural concepts, values, beliefs and social practices which are embodied lexically and idiomatically in the source-language expression.

In order to ensure consistency and transparency, all translated terms were assessed according to specific qualitative criteria. For instance, when the dimension in question was fully preserved, then a *High* rating was assigned for that translation. A *Moderate* rating also indicated that the translation of the cultural term conveys the main meaning but partly loses its cultural or contextual specificity. A *Low* score meant that there was significant loss, distortion or alteration of the meaning, communicative function or cultural significance. Following the evaluation of the three dimensions, an overall score was given to each translation. Those translations that preserved consistently across all three were categorized to be *Successful*. On the other hand, translations that were viewed to achieve partial preservation were rated as *Partially Successful* and those that did not preserve essential semantic, pragmatic or cultural features were classified as *Unsuccessful*.

Translation quality evaluation was performed through qualitative close analysis rather than utilizing automatic translation metrics. To achieve a more consistent analysis, the same criteria were systematically applied to all selected terms. The qualitative discussion that follows the assigned ratings reveals the relative strengths and weaknesses of each ChatGPT 5.2 translation. Based on the recurrent translation procedures established in the study, CSTM is proposed as a framework developed to offer practical guidance for enhancing the renditions of culturally loaded Arabic expressions into English.

Table 3. Evaluation Criteria of the Context-Sensitive Translation Model

Dimension	High	Moderate	Low
Semantic Equivalence	Preserves the complete denotative meaning.	Preserves the main meaning with minor loss of precision.	Distorts or omits the intended meaning.
Pragmatic Equivalence	Preserves the communicative function and contextual use.	Retains the general communicative intention with some contextual loss.	Fails to preserve the intended communicative function.
Cultural Equivalence	Successfully conveys the embedded cultural meaning.	Conveys only part of the cultural significance.	Omits or substantially reduces the cultural meaning.

Data Analysis

This section provides the data analysis as well as the obtained findings. It discusses the ChatGPT translation performance with respect to the proposed CSTM and illustrates how the semantic, pragmatic and cultural meanings of the cultural terms are preserved across the selected cultural categories.

Table 4. Evaluation of Chatgpt Translation of Religious and Charitable Concepts

Arabic Term	ChatGPT Translation	Semantic Equivalence	Pragmatic Equivalence	Cultural Equivalence	Final Evaluation
بركة (Barakah)	Blessing	High	Moderate	Low	Partially Successful
الوقف (Waqf)	Endowment	High	Moderate	Low	Partially Successful

Translating the cultural term بركة (*barakah*) as "blessing" presents the semantic equivalent as such a rendition reflects the general meaning of goodness, divine favor or grace. Nevertheless, this rendering does not provide pragmatic and cultural equivalence. In Arabic-Islamic usage, the term بركة *barakah* (blessing) does not just refer to a blessing in general but connotes lasting divine increase, spiritual benefit and abundance and goodness exceeding material measurement. Thus, while the rendition "blessing" is natural and understandable to English readers, it minimizes the culturally embedded meaning of this term. Consequently, this translation is regarded partially successful as revealed in Table (4) above.

Likewise, the cultural term وقف (*waqf*) is translated as "endowment" which preserves the primary semantic meaning of dedicating property or wealth for a certain purpose. However, this rendering does not fully capture the Islamic legal and religious meaning. From Arabic-Islamic perspective, the term وقف *waqf* (endowment) is the permanent dedication of an asset for religious, charitable or public benefit, with specific legal rules governing its use, ownership and continuity. The term وقف *waqf* "endowment" expresses the institutional dimension but weakens the legal, religious, cultural specificity of the original notion. Therefore, this translation could be assessed as partially successful.

As shown in Table (4) above, both examples demonstrate how ChatGPT can generate lexically acceptable translations, but it generalizes culturally loaded meanings. This finding shows the need for CSTM for assessing semantic accuracy and pragmatic function and cultural depth.

Table 5. Evaluation of Chatgpt Translation of Islamic Legal Concepts

Arabic Term	ChatGPT Translation	Semantic Equivalence	Pragmatic Equivalence	Cultural Equivalence	Final Evaluation
عِدَّة (`Iddah)	Waiting period	High	Moderate	Low	Partially Successful
المحرم (Mahram)	Guardian	Moderate	Low	Low	Unsuccessful

Translating عِدَّة (*'iddah*) as "waiting period" shows a high level of semantic equivalence, as this translation conveys the basic meaning of the obligatory period observed after the death of a spouse or divorce. Nevertheless, its pragmatic equivalence is only moderate in that the English expression does not conveying the essence of the legal aspect and religious significance that this period serves within Islamic family law. From a cultural perspective, this rendering is less accurate because it omits the broad Islamic framework governing عِدَّة *'iddah* (waiting period) such as its legal regulations, religious significance and social purpose. Therefore, although the English equivalent of عِدَّة *'iddah* i.e. (waiting period) perfectly reflects the essential

lexical meaning of this term, it does not capture its full cultural dimension. Therefore, such a rendition is regarded as partially successful.

Nonetheless, translating the cultural term *محرم mahram* as "guardian" could be considered less precise. While the English term i.e. "guardian" shares some semantic features with the Arabic term *mahram* in some contexts, it does not reflect its principal meaning. In Islamic law, the Arabic cultural term *محرم mahram* denotes any male relevant whom a woman is permitted to marry due to either blood relationship, breastfeeding or marriage ties. However, the English term "guardian" refers to any person legally responsible for another and does not carry the religious and kinship-based meanings which are linked to the Arabic term *محرم mahram*. Accordingly, it can be stated that the translation of *محرم mahram* as "guardian" shows a moderate degree of semantic equivalence, but it does not preserve the cultural specificity as well as the pragmatic function of the original Arabic term. This rendition, i.e. guardian, is, therefore, considered unsuccessful.

The obtained findings reveal that ChatGPT performs well while translating Islamic and Arabic legal concepts which have lexical equivalents such as the lexical term *عِدَّة ʿiddah* (waiting period) compared to those which represent unique legal and cultural institutions like *محرم mahram*. The results of the study also show that lexical similarity is not solely sufficient for translating cultural terms. The study also shows that, as proposed by CSTM, semantic, pragmatic and cultural dimensions should be used for assessing the AI translation of culturally loaded expressions.

Table 6. Evaluation of Chatgpt Translation of Governance and Institutional Concepts

Arabic Term	ChatGPT Translation	Semantic Equivalence	Pragmatic Equivalence	Cultural Equivalence	Final Evaluation
الشورى (<i>Shura</i>)	Consultation	High	Moderate	Moderate	Partially Successful
الديوان (<i>Diwan</i>)	Council	Moderate	Moderate	Low	Partially Successful

The data analysis shows that ChatGPT provided a translation for the cultural term *شورى shura* as "consultation" which could be considered less accurate; this is because this rendition demonstrates a high level of semantic equivalence only. It just communicates the essential meaning of this Arabic cultural term, i.e. consulting others or seeking advice before making a decision. Nonetheless, this English translation i.e. "consultations" achieves a moderate level of cultural and pragmatic equivalence. The cultural term *شورى shura* in Islamic political thought does not just refers to an expression of consultation; its essence lies in a fundamental principle of governance as well as collective decision-making established in the Islamic political tradition and Qur'an. While the English term "consultation" conveys the main communicative function but does not fully reflect the institutional, ethical and religious aspects associated with *شورى shura*. Thus, it translates the core lexical meaning of the term while retaining its broader cultural significance; this translation is, therefore, considered partially successful.

The rendering of *ديوان diwan* as "council" does not reflect the exact meaning of the Arabic term. Even though *council* conveys a potential function of a *diwan*, it presupposes only one of the many potential roles undertaken by a *diwan* in terms of administrative or consultative function. Historically, *ديوان diwan* (council), for example, has denoted a number of administrative institutions, government departments, official registers and even royal or tribal reception halls at different times during Arab and Islamic history. Therefore, the English equivalent restricts the semantic scope of the source term and does not incorporate its institutional and historical depth. This restriction is evidenced in the relatively moderate ratings for pragmatic and semantic equivalence as well as the low rating for cultural equivalence. As a result, the translation is partially successful because it just communicates one contextual interpretation and misses the broader historical and cultural meanings embedded in the Arabic concept.

The obtained findings show that ChatGPT reasonably succeeds in translating governance-related concepts which have more comparable English equivalents. However, it favors concepts with a rich institutional and historical context by choosing general lexical alternatives. This pattern highlights the inadequacies of dependence on using lexical correspondence as well as the significance of integrating contextual and cultural considerations when translating governance-related terms.

Table 7. Evaluation of Chatgpt Translation of Hospitality Points Concepts

Arabic Term	ChatGPT Translation	Semantic Equivalence	Pragmatic Equivalence	Cultural Equivalence	Final Evaluation
المجلس (Majlis)	Sitting room	Moderate	Moderate	Low	Partially Successful
الضيافة (Diyafah)	Hospitality	High	High	Moderate	Partially Successful

Translating مجلس (*majlis*) by ChatGPT as "sitting room" demonstrates some degree of semantic equivalence by reflecting the physical aspect of term. This rendition, however, ignores its social function. The cultural term مجلس *majlis* (sitting room) in several Arab societies does not just denote a sitting room; it is a cultural institution where guests, community leaders or family members gather for communal or social conversation, conflict resolution and reconstruction of social ties. Although this rendering, i.e. sitting room, is functional, it does not reflect the cultural significance and broader communicative function of the original term within its domestic contexts. Accordingly, cultural and pragmatic equivalence are both restricted and thus this translation is considered partially successful.

On the other hand, the rendition of ضيافة *diyafah* as "hospitality" presents a high level of equivalence. From semantic perspective, this translation reflects the core meaning of welcoming and hosting guests. Besides, it provides high pragmatic equivalence in that such rendering performs the same communicative function in English through referring to the receiving and treating guests with respect and generosity. Furthermore, the concept of ضيافة *diyafah* in Arab culture goes beyond the common idea of hospitality to represent social and moral obligations grounded in Arab traditions and values, including generosity, honour and protection of family or tribal honour. These cultural associations are partially reflected in the English equivalent i.e. "hospitality". Thus, this rendering undeniably conveys the primary meaning and intended purpose; however, it fails to preserve fully the broader cultural values entailed in the Arabic concept, and thus it is assessed as partially successful.

These results demonstrate that ChatGPT is good at translating concepts with relatively close lexical English equivalents, such as ضيافة *diyafah* (hospitality) than those whose meaning is strongly determined by social institutional and cultural practices such as مجلس *majlis* (sitting room). The data analysis also reveals that lexical equivalence does not ensure cultural fidelity. The findings thus reinforce the argument of this study that culturally loaded expressions should be evaluated not only on a straight semantic level, but rather in light of the semantic, pragmatic and cultural dimensions as proposed by CSTM.

Table 8. Evaluation of Chatgpt Translation of Family Values

Arabic Term	ChatGPT Translation	Semantic Equivalence	Pragmatic Equivalence	Cultural Equivalence	Final Evaluation
صلة الرحم (Silat al-Rahim)	Maintaining family ties	High	High	Moderate	Partially Successful
القوامة (Qiwamah)	Guardianship	Moderate	Moderate	Low	Partially Successful

The rendition of صلة الرحم (*silat al-rahim*) as "maintaining family ties" presents a high level of both semantic and pragmatic equivalence. This English rendition of صلة الرحم (*silat al-rahim*) precisely conveys the essential concept of maintain relationships with relatives, retaining an equivalent communicative function of

meaningful communication through highlighting value of sustaining family ties for example. Nonetheless, cultural equivalence is only moderate in that such rendition does not fully reflect the wider moral and religious value of صلة الرحم *silat al-rahim* (maintaining family ties) in Arabic and Islamic culture where this concept goes beyond just keeping in touch with relatives but constitutes a religious duty which entails support and compassion for one another, constant contact and upholding family ties. These cultural and spiritual dimensions are not fully reflected in the English translation. The translation is, therefore, deemed partially successful.

The translation of the term قوامة *qiwamah* by ChatGPT as "guardianship" provides moderate pragmatic and semantic equivalence as well as a low level of cultural equivalence. This English term "guardianship" denotes a set of responsibilities, rights, protection and financial maintenance within the family rather than just authority or legal guardianship. Therefore, English equivalent restricts the concept of قوامة *qiwamah* to the general meaning, ignoring its religious, legal and social dimensions. Furthermore, this translation of قوامة *qiwamah* as "guardianship" might mislead English speaking readers and associate the term with parental guardianship or legal custody which differ from its actual meaning in Islamic law. Therefore, the semantic content in such rendition is partially preserved since the cultural and pragmatic features of the source expression are not properly reflected in this translation.

Overall, the results of this study indicate that ChatGPT makes moderate progress on successfully translating family-linked concepts that have equivalent English terms such as صلة الرحم *silat al-rahim* (maintaining family ties). On the contrary, it does not perform well when used to translate expressions originating from Islamic law and social systems where the meanings go beyond lexically equivalence to include cultural values, religious principles and institutional practices. To this end, such findings support the core underlying assumption of this study that translational validity of culturally loaded expressions can only be achieved when linguistic aspects (i.e. meaning and context) and the cultural framework are considered together, as proposed by CSTM.

Table 9. Evaluation of Chatgpt Translation of Social Norms

Arabic Term	ChatGPT Translation	Semantic Equivalence	Pragmatic Equivalence	Cultural Equivalence	Final Evaluation
العيب (Al-'Ayb)	Shame	Moderate	Moderate	Low	Partially Successful
الفرعة (Faz'ah)	Helping others	Moderate	Moderate	Low	Partially Successful

The rendition of العيب (*al-'ayb*) as "shame" only shows a moderate semantic and pragmatic equivalence. Nevertheless, it presents low cultural equivalence. For instance, the English term "shame" refers to the general idea of embarrassment or social disapproval but does not express what العيب *al-'ayb* means within Arab culture. The Arabic idea goes beyond affective meaning and functions as a social norm governing the relationship between acceptable behaviour and community values, customs, and expectations. It is employed frequently as an implicit mechanism of social control, identifying behaviour which could not be considered as culturally appropriate even when it is neither religiously forbidden nor prohibited. Consequently, although the rendition of العيب *al-'ayb* as "shame" reflects part of the intended meaning, it does not retain the normative and culturally embedded nature of the Arabic term i.e. العيب (*al-'ayb*). Thus, this English rendering of العيب (*al-'ayb*) is regarded as partially successful.

Similarly, the rendition of الفرعة (*faz'ah*) as "helping others" just achieves moderate semantic and pragmatic equivalence. Yet, this English rendering provides a low level of cultural equivalence. In this sense, the English phrase conveys the general action of providing assistance and consequently maintains the basic communicative intention of the original term. Nevertheless, the act of فرعة *faz'ah* (helping others) represents a much cultural practice across many Arab societies. It refers to the immediate and collective mobilisation of family members, neighbours or people from the community to assist a person or household in trouble; this act portrays old culture values of solidarity, loyalty and social responsibility. The translation

of *فزع* *faz'ab* as (helping others) overlooks its collective and culturally specific dimensions. Thus, the translation preserves the main meaning but does not convey its general social significance. Thus, this translation is regarded as partially successful.

These results indicate that ChatGPT struggles to translate concepts which primarily serve as social norms rather than institutions or physical objects. Although the ChatGPT provides reasonably accurate representations of general meanings for some terms, for example, *العييب* *al-'ayb* (shame), *فزع* *faz'ab* (helping others), it reduces culturally embedded concepts to broad lexical equivalents; this overlooks the social values and communal practices that show their contextual meanings in Arab culture. This pattern also supports the notion that effective translation of culturally loaded terms needs to be given due care during translation, not only in terms of semantic meaning but their pragmatic and semantic meanings besides cultural contexts as proposed by the CSTM.

Table 10. Evaluation of Chatgpt Translation of Tribal and Honour-Related Concepts

Arabic Term	ChatGPT Translation	Semantic Equivalence	Pragmatic Equivalence	Cultural Equivalence	Final Evaluation
النخوة (Nakhwa)	Chivalry	Moderate	Moderate	Low	Partially Successful
الثأر (Tha'r)	Revenge	High	Moderate	Low	Partially Successful

The analysis shows that the cultural term *نخوة* (*nakhwa*) is translated by ChatGPT as "chivalry" which shows moderate semantic and pragmatic equivalence. However, this rendition demonstrates low cultural equivalence. From a semantic perspective, the English term "chivalry" expresses courage, honour and noble conduct, which is somewhat included by the Arabic concept *نخوة* (*nakhwa*). Furthermore, this English word pragmatically reflects some intended social values which are associated with honorable behaviour. This rendition does not, however, capture the broader cultural significance of the Arabic term in that the meaning of this concept in Arabic goes further from personal honour to loyalty, solidarity and generosity, and moral obligation to assist relatives, neighbours and members of the community especially when they are in distress. The English term "chivalry" fails to capture these culturally embedded dimensions, and thus this translation is assessed to be partially successful.

Translating *ثأر* *tha'r* by ChatGPT as "revenge" shows a semantic equivalence as this English translation communicates the essential concept of retaliating for a perceived injury or wrong. Nevertheless, this rendering, i.e. "revenge", however, shows only moderate pragmatic equivalence and very low cultural equivalence. In fact, the Arabic term of *ثأر* *tha'r* does not just refer to an act of personal revenge but a historically and socially embedded practice linked to collective responsibility, tribal honour, and customary legal traditions. It typically encompasses responsibilities that extend beyond the individual to the extended family or tribe and is regulated by rules rooted in old social customs. However, the English translation "revenge" captures an idea of retaliation but falls short of the notion's collective, historical and cultural associations. Therefore, while the translation "revenge" captures the essence of its semantic content, it fails to reflect the communicative and cultural dimensions of the Arabic term *ثأر* *tha'r*. Thus, the English translation of the Arabic term is evaluated as partially successful.

Overall, the obtained results reveal that ChatGPT presents translations of tribal and honour-related concepts through choosing general equivalents and retaining the core meanings of these cultural terms. Nevertheless, it overlooks the social structures and cultural traditions that lend significance to these notions. In addition, this pattern provides more evidence for the current study to argue that culturally loaded expressions should be assessed based on more than lexical correspondence; this highlights the value of the proposed CSTM in highlighting the cultural meanings lost during translation.

Table 11. Evaluation of Chatgpt Translation of Social Relationship Concepts

Arabic Term	ChatGPT Translation	Semantic Equivalence	Pragmatic Equivalence	Cultural Equivalence	Final Evaluation
الواسطة (Wasta)	Connections	Moderate	Moderate	Low	Partially Successful
التكافل (Takāful)	Solidarity	High	Moderate	Moderate	Partially Successful

Translating واسطة (*wasta*) as "connections" also reflects a low degree of cultural equivalence as well as moderate semantic and pragmatic equivalence. From a semantic perspective, the English term "connections" does convey the idea of relying on personal networks to access services or opportunities, reflecting partially the meaning of Arabic concept واسطة (*wasta*). From a pragmatic perspective, the intended meaning of واسطة (*wasta*) is broadly captured in the translation as this rendition communicates the essential communicative function of using interpersonal networks. Nonetheless, this rendering does not capture the broader cultural significance of the Arabic term as this concept i.e. واسطة (*wasta*) encompasses informal mediation, reciprocal social obligation, family influence and established patterns of interpersonal trust. Depending on the context in which it is used, واسطة *wasta* might have either positive or negative connotation, extending well beyond the English concept "connections". Thus, while this English translation accurately captures the general idea of the Arabic term, it does not preserve its cultural complexity, and therefore this translation is considered partially successful.

Translating تكافل (*takāful*) as "solidarity" by ChatGPT would yield a significant level of semantic equivalence, as it conveys the essential idea of mutual support and cooperation. However, the offered translation i.e. (solidarity) falls short of pragmatic and cultural equivalence. In Arab and Islamic contexts, the term تكافل (*takāful*) (solidarity) goes beyond the Muslim notion of community and refers to a more extensive systems of mutual responsibility in which individuals and communities share collective obligations supporting one another socially, morally, and in some cases, financially. The English translation i.e. "solidarity" captures the overall principle of mutual support and unity, but it does not show the ethical, institutional and religious dimensions linked with the Arabic cultural term تكافل (*takāful*). Consequently, the English translation "solidarity" is deemed partly successful.

To conclude, the obtained findings suggest that ChatGPT reasonably performs while translating concepts associated with social relationships by expressing their essential meanings through familiar English expressions. However, it often simplifies culturally embedded notions in which their meanings are formed by social norms, ethical obligations and long-established community practices. This pattern enhances the significance of taking cultural context into account besides pragmatic and semantic meanings when making assessment of AI translation; this supports the CSTM applicability in highlighting areas where cultural specificity is diminished during translation.

Discussion of the Findings

The results of the current study reveal that ChatGPT generally succeeds in translating the essential meanings of Arabic cultural terms into English. This is because the semantic equivalence scored higher than both cultural and pragmatic equivalence throughout the selected data. This means that this AI system is able to provide suitable lexical equivalents of many Arabic expressions. Nevertheless, presenting the lexical meanings alone by this AI system does not ensure adequacy of translation. As revealed in the analysis, several translations are deemed partially successful in that they risk losing salient elements of contextual and cultural meaning inherent in the original Arabic expressions. Only one rendition offered by ChatGPT i.e. محرم *mahram* "guardian" is considered unsuccessful. This is because this English rendition provided a

different legal concept and changed the intended meaning considerably. It is shown that the main challenges encountered by ChatGPT is not related to how to recognize words but how to understand cultural knowledge and contextual meaning associated with culturally loaded expressions.

The data analysis reveals a discrepancy between cultural representation and semantic accuracy. It is shown that ChatGPT used English equivalents which conveyed the general denotative meanings when translating several cultural expressions like *بركة* *barakah* (blessing), *شورى* *shura* (consultation), *وقف* *waqf* (endowment), *صلة الرحم* *silat al-Rahim* (maintaining family ties), *ثأر* *tha'r* (revenge) and *تكافل* *takaful* (solidarity). For instance, the rendition of *بركة* *barakah* as "blessing" conveys the concept of divine favour but this English translation does not reflect the broader Islamic perception of constant spiritual and material abundance. Similarly, the Arabic cultural term *وقف* *waqf* is rendered as endowment, reflecting its institutional function. Nonetheless, this rendition overlooks its unique charitable and legal role within Islamic tradition. Such examples show that presenting lexical equivalence alone does not represent culturally loaded concepts.

The obtained findings of the current study are congruent with those found in previous studies on AI translation. Examples of these studies include AlAfnan and Alshakhs (2025) who pointed out that AI systems are able to produce fluent translations of Arabic poetry. However, such systems do not succeed in maintaining cultural symbolism and emotional depth. Besides, Hanan and Sukayna (2025) showed that ChatGPT handled literary texts coherently but encountered challenges in translating religious references and culturally nuanced expressions. The present study findings are consistent with these observations though it adopts a different corpus of culturally loaded expressions associated with institutional, religious, legal, social and tribal domains. This suggests that the limitations discussed in previous studies are not limited to just literary texts but extend to a broader range of culturally loaded expressions.

The results also support the findings obtained in some previous studies such as (Ennouari and Houssaini, 2023; Al-Jarf, 2025) who concluded that the AI translation systems often tend to use direct lexical correspondences while translating culturally specific language. Similar tendencies were evident in translations such as *العيب* *Al-'Ayb* rendered as *shame*, *فزة* *faz'ah* translated as *helping others*, and *واسطة* *wasta* translated as *connections*. Although such English equivalents communicate the general meaning of the original cultural terms, they simplify the concepts whose significance derives from shared cultural values, social norms and collective practices. The obtained findings are different from previous studies that primarily identified such limitations in that the present study explains them systematically through the proposed CSTM by simultaneously assessing semantic, pragmatic and cultural equivalence.

In addition, the data analysis shows that ChatGPT performs more successfully when Arabic concepts have relatively close functional equivalents in English. Arabic cultural terms including *ضيافة* *diyafah* translated as *hospitality* and *صلة الرحم* *silat al-Rahim* translated as *maintaining family ties* preserved both the principal meaning as well as communicative function of the original Arabic terms. However, even these translations did not fully convey the religious, ethical and social obligations associated with the original concepts. Greater challenges faced ChatGPT when translating cultural concepts which are deeply rooted in Islamic legal or socio-cultural systems including *محرم* *mahram* (guardian) and *قوامه* *qiwamah* (guardianship); it is shown that lexical substitution leads to conceptual reduction as well as cultural loss. These results show that the translation quality is influenced by the cultural complexity as well as linguistic similarity of the Arabic cultural expressions.

An important contribution presented by this study lies in its multidimensional assessment of AI translations. The proposed CSTM differentiates among semantic, pragmatic and cultural performance rather than assessing translations solely as correct or incorrect. Such an approach shows that several translations could be linguistically acceptable while they do not preserve the cultural meanings. Finally, the obtained findings reveal that successful translation of cultural terms require more than lexical accuracy. Although ChatGPT consistently produces fluent English translations, this AI system simplifies cultural concepts, overlooking their cultural and contextual dimensions. Consequently, the proposed CSTM in the current study offers a practical framework which assists in identifying such limitations and provides a more comprehensive approach for the assessment and enhancement of AI translation of cultural expressions.

Conclusion and Recommendations

This study examined ChatGPT's performance in translating culturally loaded Arabic expressions into English through the lenses of semantic, pragmatic and cultural equivalence. Overall, the findings suggest that ChatGPT is capable of producing fluent and linguistically coherent translations, with most outputs successfully conveying the central lexical meaning of the source expressions. Nevertheless, lexical accuracy did not consistently translate into cultural fidelity. In many cases, the translations only partially reflected the religious, legal, historical and social associations that give these expressions their significance within Arab and Islamic culture. The analysis therefore indicates that evaluating AI-generated translations solely on linguistic correctness provides an incomplete picture of translation quality. When culturally embedded language is involved, the preservation of contextual meaning is equally essential.

A further contribution of this study is the development of the Context-Sensitive Translation Model (CSTM), which brings together established concepts of semantic, pragmatic and cultural equivalence within a single evaluative framework. Rather than treating these dimensions independently, the model enables a more comprehensive assessment of how meaning is maintained or altered during AI-assisted translation. Applying the CSTM to the present dataset demonstrated its value in identifying not only lexical accuracy but also areas where communicative intent and cultural meaning were weakened or lost. As such, the model has potential applications for translation researchers, professional translators, and AI developers seeking more nuanced methods of evaluating culturally specific translation. Although the results obtained in this study are encouraging, further validation using larger datasets, additional language pairs, and alternative AI translation systems would strengthen the model's broader applicability.

The findings also point to several practical implications. First, AI-generated translations of culturally sensitive texts should continue to be reviewed by human translators, particularly when the source material involves religious, legal, or culturally embedded concepts. Second, future AI translation systems would benefit from stronger mechanisms for incorporating cultural knowledge and contextual reasoning rather than relying primarily on lexical correspondence. Third, translation quality assessment should adopt a multidimensional perspective that evaluates pragmatic and cultural appropriateness alongside semantic accuracy. Finally, future research should extend the application of the CSTM to more diverse corpora, authentic communicative settings, and different AI translation platforms in order to examine its robustness across genres, language combinations, and cultural contexts.

References

- Aixelá, J. F. (1996). Culture-specific items in translation. In R. Álvarez & M. C. A. Vidal (Eds.), *Translation, power, subversion* (pp. 52–78). *Multilingual Matters*.
- Al-Khreshheh, M. (2025). A back translation analysis of AI-generated Arabic-English texts using ChatGPT: Exploring accuracy and meaning retention. *Dragoman J. Transl. Stud.*, 17, 97–117.
- Al Maaytah, S. A. (2026). Evaluating Three Neural Machine Translation Platforms for English-Arabic Translation: A Comparative Study of Linguistic Accuracy and Cultural Fidelity. *World*, 16(2).
- AlAfnan, M. A. (2025). Artificial intelligence and language: Bridging Arabic and English with technology. *Journal of Ecohumanism*, 4(1), 240–256.
- AlAfnan, M. A., & Alshakhs, T. (2025). Bridging Linguistic and Cultural Nuances: A Comparative Study of Human and AI Translations of Arabic Dialect Poetry. *Advances in Artificial Intelligence and Machine Learning*. 2025; 5 (1): 186. *Computational Linguistics*. <https://doi.org/10.54364/AAIML>.
- AlGhamdi, A. M. A., & Ahmed Al-Sofi, B. B. M. (2025). Artificial Intelligence and the Translation of Culture-Specific Terms on Arabic Monolingual Signs at Aseer Tourist Sites, Saudi Arabia. *UB Journal for Humanities*, 4(2), 21.
- Al-Jarf, R. (2025). Translation of Arabic expressions of impossibility by AI and student-translators: A comparative Study. *Journal of Computer Science and Technology Studies*, 7(8), 288–299.

- Alkhatnai, M. (2025). The role of artificial intelligence tools in mediating Sino-Arab cultural exchanges through intercultural translation. *Babel*, 71(6), 740-769.
- Al-Qinai, J. (2012). Convergence and divergence in the interpretation of Qur'anic polysemy and lexical recurrence. *Acta Orientalia Academiae Scientiarum Hungaricae*, 65(1), 71-92.
- Azizov, D. (2024). From Idioms to Algorithms: Translating Culture-Specific Expressions in AI Systems. *Iconic Research and Engineering Journals*, 7(10), 543-551.
- Baker, M. (2018). *In other words: A coursebook on translation* (3rd ed.). Routledge.
- Bowker, L. (2023). *De-mystifying translation: Introducing translation to non-translators*. Routledge.
- Castilho, S., Moorkens, J., Gaspari, F., Sennrich, R., Way, A., & Georgakopoulou, P. (2017). Is neural machine translation the new state of the art? *The Prague Bulletin of Mathematical Linguistics*, 108(1), 109–120. <https://doi.org/10.1515/pralin-2017-0013>
- Dickins, J., Herve, S., & Higgins, I. (2016). *Thinking Arabic translation: A course in translation method: Arabic to English*. Routledge.
- Ennouari, A., & Houssaini, K. (2025). Translating cultural-loaded content from Arabic to English: Striking the balance between machine translation and human expertise.
- Freitag, M., Foster, G., Grangier, D., Ratnakar, V., Tan, Q., & Macherey, W. (2021). Experts, errors, and context: A large-scale study of human evaluation for machine translation. *Transactions of the Association for Computational Linguistics*, 9, 1460-1474.
- Hanan, A. J., & Sukayna, A. (2025). AI vs. Human: Translating Cultural Nuances in Naguib Mahfouz's *The Cairo Trilogy*. «Technology and language»(Технологии в инфосфере), 6(4), 168-193.
- Hendy, A., Abdelrehim, M., Sharaf, A., Raunak, V., Gabr, M., Matsushita, H., ... & Awadalla, H. H. (2023). How good are GPT models at machine translation? A comprehensive evaluation. *arXiv preprint arXiv:2302.09210*.
- House, J. (2015). *Translation Quality Assessment: Past and Present*. London: Routledge. 2015
- Jiao, W., Wang, W., Huang, J. T., Wang, X., Shi, S., & Tu, Z. (2023). Is ChatGPT a good translator?. *arXiv preprint arXiv:2301.08745*.
- Koehn, P. (2020). *Neural machine translation*. Cambridge University Press.
- Mehassouel, E., & Djeha, N. (2024). Translating Culture Specific Items: A Comparative Analysis of Human and Artificial Intelligence Translations. *YMER*, 23(12):
- Newmark, P. (1988). *A textbook of translation*. Prentice Hall.
- Nida, E. A. (1964). *Toward a Science of Translating*. Leiden: E. J. Brill.
- Pedersen, J. (2011). *Subtitling norms for television: An exploration focusing on extralinguistic cultural references*. John Benjamins.
- Taher, R. S. H. (2025). The effectiveness of AI-driven translation technologies in mediating cultural understanding: A case study of English language teaching practices in Libyan higher education. *Libyan Journal of Educational Research and E-Learning (LJERE)*, 01-16.
- Vishwakarma, V. K. (2023). Translating cultural nuances: Challenges and strategies. *ELT Voices*, 13(2), 8268531.