

Article

The Shortcomings of Natural Language Processing (NLP) Models and Their Applications in Achieving Accuracy in Translating from Arabic to English

Hilal Abdul-Raziq Sadiq ^{*} , Zaxid Maxmudovich Islamov , Raziya Baltabaevna Matibaeva , Akbar Akmalldjanovich Kazakbaev  and Mutabarkhon Amanullaevna Madaeva 

Department of Arabic Language and Literature, International Academy for Islamic Studies, Tashkent 100011, Uzbekistan

* Correspondence: h.abdelrazek@iiu.uz

Received: 22 February 2026; **Revised:** 13 March 2026; **Accepted:** 1 April 2026; **Published:** 24 July 2026

Abstract: Translation is defined as the process of transferring meaning from one language to another. It is an extremely difficult and complex process because it involves not only transferring words but also ideas, culture, linguistic customs, and meanings derived from syntactic elements, their arrangement, word structure, and derivation. This is especially true in languages with complex structures, such as Arabic, which is characterized by its multiple linguistic contexts. These contexts have not been adequately addressed by NLP (Natural Language Processing) applications in machine translation models due to the lack of diverse contexts where metaphor, figurative language, grammatical inflections, morphological patterns and their connotations, and sentence structure all play pivotal roles in determining meaning. Furthermore, spoken language, with its inherent phonetic and expressive characteristics, conveys the text into broader semantic spaces. These spaces are influenced by the effect of intonation on specific syllables, the speaker's psychological state, the listener's mood, and accompanying body language, which transforms meaning into other subtle details. All of this, and more, is absent from machine translation, no matter how hard its creators try to imbue it with human emotions and feelings. This study highlights the importance of integrating in-depth linguistic analysis, contextual semantic modelling, and cultural awareness into natural language processing-based translation systems. By combining traditional linguistic insights with computational methods, the research offers a framework that can contribute to improving the accuracy of machine translation from Arabic to English.

Keywords: Language and Language Culture; Stress and Intonation; Morphological Meanings; Syntactic Analysis; Semantic Analysis; Context; Linguistic Taboos

1. Introduction

Language is a human activity practiced by a single group, characterized by social interaction and consensus. Every language contains a vast repertoire of words, expressions, structures, and grammatical forms [1]. There are other details stemming from the culture of the language itself [2], the act of speaking, and the meanings acquired from the phonological [3], morphological [4], and syntactic levels. Punctuation in written texts [5], stress [6], intonation [7], and delivery in spoken texts, along with many other important factors in the translation process. All contribute to achieving the highest degree of accuracy in the translated text.

Machine translation is a means of transferring meaning from one language to another. It is characterized by its speed and time-saving benefits for its users [8]. However, despite these advantages, it cannot be relied upon entirely and must be carefully reviewed because it often provides an equivalent for every word in the text to maintain the word count [9], which sometimes leads to a loss of the overall meaning [10]. The arrangement of words in each language depends on its system and culture, and this arrangement is not consistent at a uniform pace in all languages. Arabic, in particular, is a language with a special nature in the structure of its words and the characteristics of its sentence structure. Derivation, metaphor, context, and sounds all serve the meaning. Natural language processing and its applications in neural translation lack a large part of these subtle linguistic phenomena. Moreover, customs, traditions, values, and societal culture influence linguistic output. For example, there are linguistic taboos that cannot be used because of the social embarrassment they cause. The Arabic language euphemisms these prohibitions with other socially acceptable words. In Arabic, some body parts cannot be expressed explicitly and directly, and the original words are not used explicitly. This linguistic taboo does not exist in other languages where there is social freedom, and it is possible to express these prohibitions explicitly in other languages. Despite the tremendous progress in machine translation programs and programmers' attempts to design translation software with a human touch, the difference between machine and human translation remains clear. No matter how sophisticated, translation programs cannot provide an accurate translation of a poem [11], where the poet uses imagery from their environment to represent their symbolic meanings or express their personal feelings a fundamental element. Poetry is truly understood only by those who speak its language, and it only becomes meaningful when the reader's emotional state aligns with the writer's. Furthermore, the body language that accompanies meaning is entirely absent in text translation.

In spoken texts, machine translation programs lack the vocal expressions that can alter literal meanings and transform them into meanings related to the speaker's psychology. For example, the English verbs "went" and "go" are two different words expressing temporal values, which is the natural equivalent in translation. However, the first verb might be accompanied by a subdued tone indicating regret, while the second might be accompanied by a strong tone indicating anger [12]. All of this, and more, constitutes one of the problems that hinder the production of a complete and accurate text through neural translation applications. In the following papers, we will present the most important linguistic phenomena that are missing in neural translation programs through the translation of words, expressions, and morphological forms through Google Translate.

2. Contribution of the Study

The linguistic phenomena addressed in this study represent key aspects of the semantic shortcomings of current Natural Language Processing (NLP) models in Arabic-English translation. Unlike previous studies that have focused primarily on computational and algorithmic aspects, this research emphasizes the role of the deep linguistic structure of the Arabic language, which is closely linked to its rich cultural context.

The findings of this study provide a theoretical framework that can support researchers and developers working on improving NLP-based translation systems. By highlighting these linguistic phenomena through practical application, the study demonstrates that such features remain insufficiently captured by current systems, despite their continuous development.

3. Literature Review

The rapid evolution of machine translation (MT) has significantly enhanced cross-linguistic communication; however, achieving contextually appropriate and culturally accurate translation remains a formidable challenge. Scholars increasingly argue that translation is not merely a lexical exchange but a sophisticated transfer of ideas, cultural norms, and linguistic customs. Despite the "exciting leap forward" represented by neural machine translation (NMT), certain linguistic phenomena and structural divergences remain largely "out of reach" for modern systems. This is particularly evident in Arabic, where the intersection of structural complexity and cultural depth poses unique hurdles for Natural Language Processing (NLP).

3.1. The Human-Machine Parity Debate and Cultural Bias

A significant debate in current literature concerns whether NMT has reached parity with human translation. While some claim that intelligent systems can now translate at a human level, Das argues that NMT remains inherently inadequate because human efficiency in handling the intersection of diverse language structures and cultures is highly improbable for machines at this stage [13]. A critical issue is the inherent bias in AI design; because systems are shaped by the social experiences and worldviews of their designers, they often project a biased framework that struggles to handle non-Western linguistic constructs accurately.

In the context of Arabic, this lack of cultural “human efficiency” is a major shortcoming. Alkhatib and Shaalan note that while machine translation of Arabic has made significant progress, it still struggles to grasp the subtle semantic, cultural, and contextual nuances that characterize professional human translation [9]. Naveen and Trojovský further emphasize that current systems often fail to disambiguate polysemous words or handle the idiomatic expressions and cultural nuances inherent in the source text [14].

3.2. Morphological Complexity and Representation Gaps

The most “linguistically distinctive” aspect of Arabic is its morphology, which accounts for a substantial portion of the language’s processing difficulty. As a highly sophisticated inflectional language, Arabic utilizes a complex system of roots and patterns to derive meaning. However, Habash identifies a critical “representation gap” in MT research: while dictionaries use lexemes, many statistical and neural models rely on undiacritized word stems. This mismatch results in a loss of semantic granularity, as the system cannot distinguish between subtle grammatical functions or permanent character traits embedded in word structures [15].

Furthermore, Nzeyimana points out that standard sub-word tokenization methods are limited to the “surface forms” of words. In low-resource or morphologically rich settings, this prevents the model from understanding underlying linguistic information. Explicit morphological modeling and the use of “lexeme-and-feature” levels of representation are necessary to ensure that the system recognizes the full scope of Arabic’s derivational richness [4].

3.3. User Perspectives and Syntactic Challenges

Research into the practical application of translation tools highlights the ongoing reliance on digital translation platforms (DTPs) despite these known shortcomings. Alamri and Al-Amri found that academic faculty members frequently consult websites like Google Translate and smartphone apps, even though satisfaction with translation quality varies and there are no statistically significant differences in perceived satisfaction across different user groups. This suggests a persistent need for improvement in MT systems to meet the rigorous standards of academic and abstract translation [16].

These improvements must address Arabic’s flexible syntax—characterized by the ability to delay subjects or reorder objects for emphasis—which presents a structural divergence that rigid English-based models struggle to bridge. This complexity is even more pronounced in spoken language, where Arabic natural language generation (NLG) is currently underdeveloped. Abdel Monem et al. highlight that in spoken dialogue, the “intent” of the speaker is paramount. Traditional tools are often insufficient because they fail to account for the interplay between syntax and the speaker’s psychological state [17].

3.4. Research Gap

Although all these studies added valuable insights regarding the challenges facing Arabic in machine translation to and from English, they overlooked some important points and details revealed by this research. Some studies that focused on the phonetic aspect lacked the significance of sound characteristics such as voicing or emphasis as indicators of strength, as opposed to voicelessness or attenuation as indicators of weakness. These studies focused more on the formal aspect related to the presence of vowel sounds and their morphological changes [17]. Similarly, studies that discussed the morphological aspect focused more on morphemes and their meanings, such as prefixes and suffixes [4,5], and did not pay attention to the morphological meaning of the form, whether in the structure of verbs or derived adjectives. Those who focused on the syntactic aspect focused on personal pronouns, gender, singular, dual, and plural [16], without mentioning the importance of inflection in determining and directing meaning. None of the studies that addressed the cultural context mentioned body language or linguistic taboos in Arabic

culture. Therefore, this study differed in a number of important details that should be taken into account. Consideration should be given when designing neural translation applications, especially since the participants in this research are specialists in linguistic sciences, particularly the Arabic language.

4. Methodology

This research adopts a descriptive-analytical linguistic approach, evaluating the translation output produced by Google Translate as a model neural machine translation system. The analysis focuses on key linguistic dimensions, including morphological patterns, inflection, word order flexibility, idiomatic expressions, and culturally embedded meanings to demonstrate the shortcomings of machine translation in Arabic.

The references used in this research consist of a collection of studies that discussed machine translation and its evaluations, in addition to several studies that addressed the relationship between society, culture, emotions, and language, and their impact on translation.

5. Results and Discussion

- The applications of natural language processing and neural translation are limited for some morphological phenomena, especially the functions of verb morphological patterns.
- The inability to produce an accurate translation of sentences in which the subject is delayed and comes after the object.
- Machine translation ignores the diacritical marks that guide the meanings of words, whether at the level of the word's structure itself, or at the level of the word's function within the sentence.
- Machine translation applications lack cultural context, which includes the speakers' culture, linguistic customs, symbol meanings, colors, and body language. This deficiency leads to literal translations that distort the meaning.

Languages are divided into isolate or rigid languages, which rely on monosyllabic units. These units combine to form distinct meanings within the linguistic community for communicative purposes. These languages are often rigid, with limited content and range.

The other category of languages is agglutinative languages, whose grammatical relationships are formed by adding specific morphemes that function as adjectives or modifiers. Turkish is an example of this. Then there are inflectional or derivational languages, which are more widespread and rely on a formal structure that aims to distinguish meanings through prefixes, suffixes, and inflections. Examples include Arabic, Russian, and English. However, Arabic is considered the most sophisticated, as it is a language that relies entirely on derivation and morphemes to express its meanings.

5.1. Language and Language Culture

Learning a language involves two things: studying its structure and vocabulary, and understanding the relationship between its structures and words and the non-linguistic reality, civilization, and culture of that language which is something else entirely [18]. This is where the difficulties arise when learning a language without understanding the contexts in which its words and structures are used. Contrary to popular belief, learning a foreign language outside its native country is not considered a long and arduous process. The direct, active, audio-visual, and immersive methods are all considered effective ways to learn a language and understand its specific characteristics simultaneously.

Bilingualism and its culture are similar to bilingualism and speech. Language is the set of rules and rigid laws found in grammar books, morphology texts, and dictionaries, and it can be learned within a specific timeframe. However, the act of speaking or talking cannot be learned within the same timeframe [12]. This is because speech is an act linked to the culture of a language, which arises from the set of beliefs, values, customs, historical events, and environmental, psychological, political, and economic factors specific to a given society. These factors shape this society into a linguistic community with specific characteristics. The culture of a language is manifested in idiomatic expressions, proverbs, sayings, metaphors, symbols, signs, and quotations [19]. The culture of a language is the figurative meaning that encompasses ideas, feelings, and emotions. A machine translator should not be equipped with the rules and vocabulary of a language without its cultural context [20]. Translation is essentially the transfer of

ideas, and these ideas, which originate within a linguistic framework, are imbued with the societal culture in which the text's author was raised. If a translator (human or machine) merely transfers words without conveying the ideas, the resulting text will be flawed, with meaning lost or partially missing. How can a translator render phrases like "If water is available, dry ablution is invalid" and "You are divorced three times" without knowledge of Islamic culture? Similarly, how can they translate expressions like "black market," "money laundering," or "market Sharks" without understanding the economic conditions of the target linguistic community? Furthermore, within a language's culture, there are linguistic taboos that the language does not explicitly use but rather alludes to through metaphors. These taboos can backfire in translation, as machine systems might express these terms directly because machines are not embarrassed. Not everything that is taboo is forbidden in every culture; some languages explicitly state what is prohibited. This is where customs, traditions, and societal norms come into play [21]. Therefore, those responsible for developing translation programs must ensure they possess a deep understanding of the linguistic culture rooted in the environment, society, customs, traditions, and religious references. One of the most important examples related to culture is the Quranic text. In Arabic culture, the readings of some words may vary, but the meaning remains the same [22]. There are also many examples of language whose semantic meaning cannot be grasped or translated through literal translations alone. For example, jokes cannot be translated because they are intrinsically linked to the culture of the language [23]. A listener will not laugh at a joke about a famous person in the joke-teller's community if they do not know that person, or a joke involving a historical event or a set of customs or values they do not believe in. Even food, its varieties, and aromas are all intertwined with the culture of the language.

5.2. Morphological Values

Those responsible for machine translation programs must integrate the subtle nuances of morphological meanings into their neural translation applications. Morphological meanings constitute important values in the semantic content of the text [24]. Machine translation cannot distinguish between the added meaning in the verb form (yatafaalu **يَتَفَعَّلُ**) and the form (yafalu **يَفْعَلُ**), as will be shown in **Table 1**.

Table 1. Meanings of Verb Forms.

Morphological Form	Verbal Form	Machin Translation Output	The Meaning in Arabic Language
يَتَفَعَّلُ - يَتَفَعَّلُ Yafaelu - yatafaalu	يَتَخَيَّرُ - يَتَخَيَّرُ Yakhtaru- yatakhayaru	Chose - choses	(yatakhayaru) He takes his time choosing without rushing.
يَفْعَلُ - يَفْعَلُ Yafalu - yatafaalu	يَتَمَشَّى - يَتَمَشَّى Yamshy- yatamasha	Walk - he walks	(yatamasha) He walks leisurely and slowly, as if in contemplation.

Machine translation rendered the two forms as having the same meaning despite their differences, and in Arabic, every increase in structure corresponds to an increase in meaning.

This is not limited to verb structures alone; some adjectives derived from verbs also do not accurately reflect their degrees of intensity and exaggeration in machine translation. Arabic has specific patterns for the structure of its nouns. For example, the active participle is on the pattern **فَاعِل** (faael), and the meaning of this pattern is variable, not fixed. Exaggeration is further emphasized in another type of derived adjective, leading to the formation of forms like **فَعَال** (faaal) and **فَاعِل** (faiel) to denote exaggeration. Natural language processing applications cannot distinguish between these forms, as will be shown in **Table 2**.

Table 2. Meanings of Derivatives.

Morphological Form	Derived Adjective	Machin Translation Output	The Meaning in Arabic Language
فَاعِل - فَاعِل Faaal - Faael	كَذَّاب - كَذَّاب Kazzab- kazez	Lier- Lier	Kazzab: means that he lies frequently. Kazez: means that lying is an acquired trait and not part of his character.
فَاعِل - فَاعِل - فَاعِل Faul - Fael- Fael	حَاسِد - حَاسِد - حَاسِد Hassid- Hassud - Hased	Envious- Envious-Envious	Hassud/Hassid: A fundamental and constant characteristic that will not change. Hased: It is temporary and not permanent.

Regarding the issue of derived adjectives and their multiple forms, it is noticeable that machine translation sometimes distinguishes between these forms, and sometimes it does not provide any equivalent for the word. For example, Google Translate differentiates between (Aakel) أَكَل “eater” and (Akuol) أَكُول “Gluttonous” and provides an accurate translation for each form, but it does not provide that distinction in many other forms.

Likewise, one of the errors that machine translation makes at the morphological level is confusion in the derivation of forms [25]. Sometimes, the translation program is unable to reach the correct root of the word and translates it with a different meaning.

5.3. Grammatical Values

Undoubtedly, every language has its own grammatical system [26]. The basic grammatical patterns in Arabic consist of meaningful sentences, or complete sentences, which are nominal and verbal. The translator must follow these sentences in the text to be translated, examining their completeness, the order of their units, any irregularities in their order, and finding an equivalent in the target language. For example, English does not have nominal sentences; its sentences are verbal and begin with a subject. Moreover, delaying one of its components can distort the meaning [26]. In Arabic, it is possible to delay the subject and place the object before it without changing the meaning, which is not possible in English. Arabic also has grammatical markers that preserve meaning, as will be shown in **Table 3**, which identify the subject and object.

Table 3. Order of Grammatical Functions.

The Sentence	Machin Translation Output	The Correct Meaning
أَكَلَ الْفَأْرَ الْقَطُّ Akala alfaar Alketa	The mouse ate the cat.	The cat ate the mouse

If we translate this sentence into English, there will be no distinction between the subject and the object, especially since English does not allow the subject to be placed after the object. According to English grammar, the mouse would be the subject. Arabic also distinguishes between masculine and feminine, singular, dual, and plural [16], and has rules for numbers. The noun following the interrogative particle “kam” (كَمْ) (how many) is singular in Arabic and plural in English. Similarly, the word order of descriptive sentences is different; in Arabic, the adjective follows the noun it modifies, unlike many other languages where the noun follows the adjective. Furthermore, machine translation programs do not distinguish between exclamatory and negative constructions, as shown in **Table 4**.

Table 4. Grammatical Markers.

	Machin Translation Output	Evaluation According to Arabic Rules
مَا أَحْسَنَ زَيْدًا Ma Ahsana Zaydan	How excellent is Zayd!	It is correct
مَا أَحْسَنَ زَيْدٌ Ma Ahsana Zaydun	How good Zayd is	He did nothing good
مَا أَحْسَنَ زَيْدٍ Ma Ahsana Zayden	How handsome Zayd is!	What is good in his action?

The variation in case endings (such as the damma, fatha, and kasra) on the noun at the end of the sentence led to a change in meaning in each of the preceding sentences. However, the machine translation did not provide any additional information to indicate these meanings. Furthermore, without case endings, it would be impossible to distinguish between the negative particle “la” (لَا) indicating number and the negative particle indicating gender, as in the example shown in **Table 5**.

Likewise, grammatical markers are of great importance in grammatical topics such as addition and specification, and machine translation cannot distinguish between these grammatical functions because it does not consider grammatical markers, as shown in **Table 6**.

Table 5. Distinguishing between “La لا” for Gender and for Number.

The Sentence	Machin Translation Output	Evaluation According to Arabic Rules
لا رجلاً في الدار LA Ragulan Fi Addar	There is no man in the house.	There is no man, whether child, youth, or old man
لا رَجُلَ في الدار LA Ragulun Fi Addar	There is no man in the house.	There isn't just one man, but possibly two or more.

Table 6. Grammatical Markers and Functions.

The Sentence	Machin Translation Output	Evaluation According to Arabic Rules
اشتريت قَدَحَ ماءٍ Ishtaraytu qadaha Maaan	I bought a glass of water.	I bought a water not glass
اشتريت قَدَحَ ماءٍ Ishtaraytu qadaha Maaen	I bought a glass of water.	I bought a glass not water

5.4. Cultural and Contextual Values

Cultural context plays a crucial role in revealing meanings [14], especially with regard to idiomatic expressions, proverbs, and maxims [19]. We must also consider the role of body language, which often conveys a significant portion of the intended meaning. For example, in film translation, we find a fundamental flaw in the translation process. This flaw doesn't stem from translators' ignorance of the words being translated, but rather from the fact that most translate the text without seeing the scenes. In this case, the text may lose a considerable part of its meaning, a significant portion represented by the body language accompanying the spoken action. Gestures or body language can express a large part of the meaning. One Arab scholar emphasized this point, saying, “I cannot understand someone speaking to me in the dark” [27], (vol 2, p. 147). This is precisely what we mean. You cannot understand someone speaking to you from behind a curtain; there are hidden meanings that can only be revealed by observing their facial expressions, eye movements, and hand gestures. Body language carries-sometimes- more than 50% of the intended meaning [28], (p. 43), and machine translation lacks this visual aspect. I am referring here to Google Translate and other translation systems that do not take body language into account in written or spoken texts, especially since body language varies from one language group to another. This is shown in **Table 7**.

Table 7. Body Language.

Body Language	Machin Translation Output	The Meaning of Body Movements in the Context of the Arabic Language
طأطأ رأسه	He lowered his head.	Shame or insecurity
ارفع رأسك عاليا	Raise your head high	Pride and confidence
زم شفطيه	He pursed his lips.	Confusion or anger
قطب حاجبيه	He furrowed his brow.	Anger and dissatisfaction
اتسع بؤبؤ عينه	His pupils dilated.	Fear or surprise
عض على أظفار مله	He bit his fingertips.	Regret

Machine translation of such body language expressions often provides a literal translation divorced from the cultural context of the language. This is particularly evident in Arabic, and therefore the output fails to convey the intended meaning as it relates to the speakers' culture.

Furthermore, a semantic shortcoming of machine translation is its failure to accurately translate the symbols and colors used within the cultural context of Arabic [29], as illustrated in **Table 8**.

Table 8. Idiomatic Expressions.

Symbolic Expression	Machin Translation Output	The Meaning of Colors in the Cultural Context of the Arabic Language
اسودَّ وجهه Iswada waghuhu	His face turned black.	Sadness
ابيض وجهه Ibyada waghuhu	His face turned white.	Joy

Table 8. Cont.

Symbolic Expression	Machin Translation Output	The Meaning of Colors in the Cultural Context of the Arabic Language
اصفر وجهه Isfara waghuhu	His face turned yellow.	Fear or illness
الضوء الأخضر Aldoo Alkhdar	He gave him the green light.	Permission granted
هذا الموضوع خط أحمر Haza almawdoo khat ahmar	This is a red line.	This subject cannot be discussed
رفع الراية البيضاء Rafaa Alraya Albida	Raising the white flag	Surrender in any battle

Colors, in a cultural context, carry greater connotations than words. When the phrase “his face turned black” is used in any context, speakers immediately understand that something sad happened that caused the man to experience deep sorrow. The same applies to the rest of the sentence. The literal translation provided by machine translation failed to convey the intended meaning.

Similarly, machine translation did not provide accurate meanings for idiomatic expressions in Arabic [9], offering a literal meaning far removed from the figurative meaning these expressions convey within their cultural context, as shown in Table 9.

Table 9. Cultural Expressions.

Idiomatic Expressions in Arabic	Machin Translation Output	Meaning of Expressions in Arabic Language Culture
يركب الموجة Yarkabu Elmouga	He Rides The Wave	A Hypocritical Person Who Supports Those With Power And Influence.
دفن رأسه في الرمل Dafana Rasohu Fi Lremal	Burying His Head In The Sand	Fear And Avoidance Of Confrontation.
بيد من حديد Beyade Men Hadid	With An Iron Fist	Power In Governance And Decision-Making.

6. Conclusion

This research has demonstrated that while machine translation (MT) and natural language processing (NLP) applications offer significant advantages in speed and global accessibility, they remain inherently limited in their ability to achieve maximum accuracy when translating from Arabic to English. The primary shortcomings identified centre on the machine’s inability to transcend literal word-for-word exchanges to capture the multi-dimensional layers of the Arabic language.

As detailed in the results, neural translation models frequently fail to account for the sophisticated morphological derivational richness of Arabic, often ignoring diacritical marks and morphological patterns that determine specific shades of meaning, such as the difference between a permanent character trait and a temporary state. Furthermore, the flexible syntactic structure of Arabic, specifically the ability to delay subjects, remains a structural divergence that often leads to distorted translations in rigid English-based models.

Beyond technical structural hurdles, the most significant barrier is the absence of cultural and contextual intelligence. Machine translation systems lack the capacity to interpret socio-cultural nuances, linguistic taboos, and the figurative meanings embedded in idiomatic expressions and symbolic colours. Additionally, the vital role of non-verbal cues, such as body language and vocal intonation, remains entirely absent from text-based MT, leaving over 50% of intended meaning “in the dark”.

Ultimately, this study underscores that for morphologically and culturally rich languages like Arabic, machine translation must be viewed as an aid rather than a replacement for human expertise. To move toward higher accuracy, developers must prioritize the integration of “lexeme-and-feature” representations and context-aware modelling that accounts for the inextricable link between a language and the societal values of its community. Further research should continue to utilize targeted “challenge sets” to systematically identify and bridge these persistent linguistic gaps.

Author Contributions

Selecting the topic, analysing the information, conducting morphological analysis, translating, and comparing different viewpoints, H.A.-R.S.; Gathering and analyzing phonetic data and identifying the linguistic terms used and contributing to the analysis, Z.M.I.; Gathering and classifying morphological data and applying it to machine translation applications, R.B.M.; Classifying tables and analyzing grammatical data; Gathering and analyzing data related to idiomatic expressions and body language, A.A.K.; Collect examples related to the rhetorical aspect and identify the references used, M.A.M. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

All sources and references used are readily available to everyone. The books are available in libraries that hold printed copies, and also as PDFs in online libraries. Published articles are also freely accessible and easy to find.

Conflicts of Interest

The authors declare no conflict of interest.

AI Use Statement

The authors declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

1. Acı, M.; Vuran Sarı, N.; İnan Acı, Ç. Morphological and structural complexity analysis of low-resource English–Turkish language pair using neural machine translation models. *PeerJ Comput. Sci.* **2025**, *11*, e3072. [[CrossRef](#)]
2. Anani, M. *The Art of Translation*, 5th ed.; Egyptian Publishing Company–Longman: Cairo, Egypt, 2000.
3. AL-Shayeb, F. *The Impact of Phonetic Rules on the Structure of Arabic Words*, 1st ed.; Alam al-Kutub al-Hadith: Amman, Jordan, 2004.
4. Nzeyimana, A. Low-resource neural machine translation with morphological modeling. In Proceedings of the Association for Computational Linguistics: NAACL 2024, Mexico City, Mexico, June 2024; pp. 182–195. [[CrossRef](#)]
5. Habash, N. Arabic Morphological Representations for Machine Translation. In *Arabic Computational Morphology*; Springer: Dordrecht, Netherlands, 2010; pp. 263–285. [[CrossRef](#)]
6. Xu, W.; Liu, C. From grammatical to phonological metaphor: Marked tonality as semogenic remapping. *Lingua* **2026**, *331*, 104091. [[CrossRef](#)]
7. Xu, W.; Liu, C. Marked intonation as a phonological metaphor: A systemic functional approach. *J. Pragmat.* **2025**, *235*, 145–163. [[CrossRef](#)]
8. Omar, L.I.; Salih, A.A. Systematic Review of English/Arabic Machine Translation Postediting: Implications for AI Application in Translation Research and Pedagogy. *Informatics* **2024**, *11*, 23. [[CrossRef](#)]
9. Alkhatib, M.; Shaalan, K. The Key Challenges for Arabic Machine Translation. In *Intelligent Natural Language Processing: Trends and Applications*; Springer: Cham, Switzerland, 2018; pp. 139–156. [[CrossRef](#)]
10. Zemni, B.; Zitouni, M.; Bouhadiba, F.; et al. A Comparative Study of the Lexical Ambiguity of Arabic, English, and French in Natural Language Processing. *J. Intercult. Commun.* **2024**, *24*, 203–212. [[CrossRef](#)]

11. Liang, L. Discrepancies in poetic discourse: A translation process from a relevance-theoretic perspective. *Soc. Sci. Humanit. Open* **2023**, 8, 100666. [CrossRef]
12. Tamam, H. *The Arabic Language: Its Structure and Meaning*; Dar al-Thaqafa: Casablanca, Morocco, 1994.
13. Das, A.K. Neural Machine Translation (NMT): Inherent inadequacy, misrepresentation and cultural bias. *Int. J. Transl.* **2020**, 32, 115–145.
14. Naveen, P.; Trojovský, P. Overview and challenges of machine translation for contextually appropriate translations. *iScience* **2024**, 27, 110878. [CrossRef]
15. Habash, N. *Introduction to Arabic Natural Language Processing*, 1st ed.; Morgan & Claypool Publishers: Toronto, ON, Canada, 2010.
16. Alamri, H.; Al-Amri, A. Arabic-English Machine Translation (MT) Quality: Exploring Researchers' Perspectives and Challenges. In Proceedings of the 14th International Multi-Conference on Complexity, Informatics and Cybernetics: IMCIC 2023, Online, 28–31 March 2023; pp. 176–182. [CrossRef]
17. Abdel Monem, A.; Shaalan, K.; Rafea, A.; et al. Generating Arabic text in multilingual speech-to-speech machine translation framework. *Mach. Transl.* **2008**, 22, 205–258. [CrossRef]
18. Kramsch, C. *Language and Culture*, 3rd ed.; Oxford University Press: Oxford, UK, 1998.
19. Al Shaikhli, M. Problems of Machine Translation Systems in Arabic. *J. Lang. Teach. Res.* **2022**, 13, 755–762. [CrossRef]
20. Castilho, S.; Knowles, R. A survey of context in neural machine translation and its evaluation. *Nat. Lang. Process.* **2025**, 31, 986–1016. [CrossRef]
21. Al-Khalil, E.A.A.; Mohammed, I.A. Some reservations on translating taboos from English into Arabic: Problems and solutions. *Wasit J. Humanit.* **2025**, 21, 1343–1329. [CrossRef] (in Arabic)
22. Al-Haj Eid, O.A.; Mansour, A.W.; Abumahfouz, A.M.; et al. Morphological markedness in the Holy Quran. *Forum Linguist. Stud.* **2025**, 7, 9. [CrossRef]
23. Omar, A.; Gomaa, Y.A. The machine translation of literature: implications for translation pedagogy. *Int. J. Emerg. Technol. Learn.* **2020**, 15, 228. [CrossRef]
24. Sadiq, H. General principles of morphological theory. *Hermis* **2024**, 13, 131–167. [CrossRef]
25. Isabelle, P.; Cherry, C.; Foster, G. A challenge set approach to evaluating machine translation. In Proceedings of the 2017 Conference on Machine Translation Evaluation, Copenhagen, Denmark, September 2017.
26. Kazemi, A.; Toral, A.; Way, A.; et al. Syntax- and semantic-based reordering in hierarchical phrase-based statistical machine translation. *Expert Syst. Appl.* **2017**, 84, 186–199. [CrossRef]
27. Ibn Jini. *The Despicable*, 1st ed.; Dar Al-Kutub Al-Masreya: Cairo, Egypt, 1953. (in Arabic)
28. Mehrabian, A. *Silent Message*; Wadsworth Publishing Company Inc.: Belmont, CA, USA, 1972.
29. Ding, Q.; Cao, H.; Cao, Z.; et al. Cross-lingual semantic information fusion for word translation enhancement. *Inf. Fusion* **2026**, 125, 103492. [CrossRef]



Copyright © 2026 by the author(s). Published by UK Scientific Publishing Limited. This is an open access article under the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Publisher's Note: The views, opinions, and information presented in all publications are the sole responsibility of the respective authors and contributors, and do not necessarily reflect the views of UK Scientific Publishing Limited and/or its editors. UK Scientific Publishing Limited and/or its editors hereby disclaim any liability for any harm or damage to individuals or property arising from the implementation of ideas, methods, instructions, or products mentioned in the content.