Transliteration of Arabic Words/Phrase into English: An Exploration of Ambiguity Markers

Majed Abdullah Alharbi[1], & Mohammad Shariq[2]

[1] Department of English Language, College of Education, Majmaah University, Al Majma'ah, Saudi Arabia

[2] Associate Professor of Linguistics, Department of English Language and Literature, College of Languages and Humanities, Qassim University, Saudi Arabia

Correspondence: Majed Abdullah Alharbi, Department of English Language, College of Education, Majmaah University, Al Majma'ah, Saudi Arabia. E-mail: m.alufi@mu.edu.sa

Received: February 25, 2024 Accepted: April 4, 2024 Online Published: April 18, 2024

doi:10.5430/wjel.v14n4p404 URL: https://doi.org/10.5430/wjel.v14n4p404

Abstract

Transliteration is a useful process when communication involves a language pair out of which each follows a different script, such as Arabic and English. One danger posed by this process is the ambiguity between the source and target of the communication. This study analyzes the intricate process of Arabic to English transliteration and the factor that make it ambiguous. The study aims to identify, categorize, and analyze the ambiguity markers that frequently arise during the transliteration of Arabic script into the Latin alphabet. The study interviewed 6 specialist translators in the Saudi context to identify the source of difficulties they encounter and which may bring ambiguities to the readers at various language levels. Results indicated that non-standardization in translation from Arabic to English was a cause of ambiguity in transliteration. Finally, regional dialects could not be adjusted in the transliteration spectrum in Google Translate. This research contributes to the field of transliteration studies by providing a comprehensive framework for understanding and addressing ambiguity in the transliteration process.

Keywords: Ambiguity markers, Arabic words/phrase, English, exploration, transliteration, education, research work

1. Introduction

In the domain of cross-cultural communication, transliteration plays a vital role as a fundamental bridge connecting diverse linguistic landscapes (Habash et al., 2007). Transliteration, by definition, involves the process of representing the sounds of one language using the writing system of another, allowing readers to grasp the phonetic aspects of foreign words (Regmi et al., 2010). Among the most intricate transliteration challenges lies in the pairing of Arabic and English. Despite their historical significance and global reach, these languages belong to entirely separate linguistic families and writing systems. Current scholarly literature has pinpointed numerous obstacles within this undertaking, with one particularly notable hurdle being the existence of ambiguity markers. These markers introduce uncertainties into how a given Arabic word or sound can be faithfully rendered in English (Deheuvels et al., 2006). Arabic, with its intricacies like a rich phonetic inventory, context-sensitive pronunciations, and diverse regional dialects, presents formidable challenges for those involved in transliteration. Past research has delved into the factors contributing to this ambiguity, encompassing variations in phonetics, morphological subtleties, and distinctions in orthographic representation. However, while these studies have shed light on various facets of the issue, there is still a pressing need for a comprehensive exploration dedicated to understanding, categorizing, and effectively addressing these ambiguity markers.

Consequently, the current study differs from the earlier investigations (Al-Sughaiyer & Al-Kharashi, 2004; Askari, 2019; Issa, 2018), in two significant ways. The study highlights the difficulties translators encounter when translating everyday language. Firstly, it explores the challenges posed by intricate, regional, and well-known Arabic phrases. Secondly, it delves into the coping mechanisms employed by translators to render these phrases into English, while also maintaining consistency with 'standard' language within the same literary piece. By addressing these topics, researchers can gain insights into the challenges translators face when translating everyday expressions, particularly those rooted in the Arabic vernacular. Understanding these challenges could inform revisions to the typology of functional translation and cultural specificity techniques.

Furthermore, the absence of ambiguity markers, orthographic differences, morphological uniqueness, phonetic features, are only some of the challenges posed by the Arabic-English language pair in the transliteration process (Issa, 2018). Without addressing these markers and the challenges they pose, the resultant English representations of Arabic content risk being inconsistent, misinterpreted, or entirely misleading (Askari, 2019). But with greater possibilities of cross-cultural communication facilitated by technological advancements, it is of much importance to explore these pitfalls to ensure smooth, clear, and coherent communication to take place. Problems in dealing with a language pair arise when some sound(s) or letter(s) of the source language do not exist in the target language. For example, the Arabic sounds (Halimah, 2016), ع, ط , ق , are only some of the problems faced when transliterating from Arabic to English as there are no exact

matches for them in the latter. Since it is the scripts (and not the languages) that transliteration deals with, change in language is not essential, in other words, for the purpose of this study, Arabic can be written in the Roman script (Masmoudi et al., 2019). This study poses the following questions:

Research questions

1. What are the primary sources of ambiguity in the transliteration of Arabic words/phrases, into English?
2. How do phonetic variations in Arabic dialects contribute to transliteration ambiguity into English?
3. What role does lexical homophony play in introducing transliteration uncertainties?

2. Literature Review

The fundamental building blocks of any language or culture are phraseological units. Acknowledging and interpreting these units correctly indicates that the translation of the expression or artwork is accurate (Tomih, 2023). One hot topic in translatology right now is the translation of phraseological units. Numerous investigations have been conducted to investigate the process of translating Arabic phraseological units into English (Hassan, 2019). Nonetheless, there hasn't been enough focus on routine phrase analysis in the academic communities of Arabic and English. Routine phrases are phraseological units with essential psychological and social functions that merge certain communicative contexts with a lack of textual independence (Rees, 2018).

Transliteration can be classified into four main types. Specific transliteration involves the conversion of text for particular purposes, where maintaining accuracy and complete faithfulness to the syllables of the source language is crucial. Partial transliteration, on the other hand, is utilized when only select utterances need to be converted, rather than the entire script. Occasional transliteration focuses on transliterating proper nouns, quotations, or specific graphemes from the source language into the target language. This type of transliteration is the most common and forms the core subject of this study. Lastly, assimilated transliteration occurs when loanwords or blends from the source language are written in the script of the target language.

The challenge of transliteration between Arabic and English has long been recognized in linguistic research. Al-Salman (2008) emphasized the phonetic variations inherent in Arabic dialects, which often lead to multiple possible English representations for a single Arabic sound or word. The differences are the outcome of the uniqueness of each of these languages and the sounds that they are composed of. Similarly, Al-Sughaiyer and Al-Kharashi (2004) explored the morphological complexities of Arabic, where word formation and the addition of affixes add to the ambiguity during transliteration. Orthography is another dimension of the problem, as highlighted by El Kholy & Habas (2012) pointing out that the Arabic writing system is characterized by its abjad nature and the absence of short vowels in standard writing, often necessitates context-based decisions during transliteration, leading to potential inconsistencies and possible ambiguities.

Another dimension to consider is the practical application of transliteration tools. As discussed by Al-Muhtaseb and Mellish (1998) stated that current computational models for Arabic-English transliteration often lack the nuance and adaptability required to handle real-world texts, especially when juggling between formal Modern Standard Arabic and regional dialects. This underscores the need for advanced models that are both robust and adaptable, equipped with a comprehensive understanding of ambiguity markers and their implications.

Ambiguity in some syntactical patterns within the same dialect has always been problematic for automated translation systems as well as translators who are humans (Prior et al., 2011). When it comes to machine translation, these difficulties multiply when dealing with anatomically dissimilar dialects like Arabic, English, and French. When the meaning nuances of Arabic words diverge greatly from those of their English counterparts, the lexical ambiguity of Arabic in Natural Language Processing (NLP) presents extra challenges (Gugliotta, 2022). When two or more Arabic words translate to a single English word, this is frequently the case. Ambiguity in NLP is also caused by semantic gaps between the two languages. Zemni et al. (2024) investigated cases of imprecision in machine translation between Arabic and English, as well as between Arabic and French. Separation, determination/non-determination, teamwork, and the role of vocabulary as useful units are among the topics covered. Additionally, the research contrasts the categories of syntax into which elements are divided in Arabic, English, and French.

Jabak (2023) examined the mistakes made in translating English news articles into Arabic by a sample of Omani undergraduate students. A unique data-gathering technique was used in the study, a brief interpretation assessment consisting of fifteen English news articles. In the second semester of the 2022 academic year, 45 third-year students at the University of Nizwa's Department of Foreign Languages were given the test. Analyzing the data revealed that most of the mistakes made by the students were in the areas of lexical and grammatical errors, acronyms, cities, and correct adverbs. The pupils' lack of familiarity with journalistic jargon, syntax, acronyms, transliteration of English metropolitan areas, and correct adverbs was the reason for these mistakes. Given the severity of these flaws, the researchers recommended conducting additional large-scale research to support, refute, or illuminate additional facets of the study's subject matter.

Al-Jarf (2022a) determined what transliteration anomalies are created when native Arabic speakers transliterate personal names with geminates to English on social media. To determine the percentage of Arabic names in which geminates were correctly transliterated into double consonants, the percentage of Arabic names in which geminated consonants were reduced to a singleton consonant in the English transliteration, and the percentage of Arabic names where a singleton consonant was doubled in the English transliteration, a sample of 406 English transliterations of Arabic personal names with geminates by Arabic native speakers was compiled from Facebook.

Three-quarters of the Arabic name tokens containing geminates were found to have been transliterated accurately, for example, the English transcription of an Arabic geminated consonant, such as in compound names like Abdullah and Noureddin, and Nassar, Algammal, Alqattan, Allam, Hagga, and son, was represented by a double consonant. In the matching English transliteration, the geminate was represented by a single consonant in 41% of the name tokens, such as Amouna, Amool, Elzahar, Hamam, Elnagar, Sedeek, and Fatouh. A single phoneme was doubled in 26% of the English transliterations to match the Arabic name, despite the Arabic name having no geminates and the consonant being pronounced as a single phoneme in Ahmmed, Anass, Ossama, Quassem, Sammar, Wissam, Yassin, and Youniss. The s, which made up 23% of the tokens, was the most frequently geminated consonant. Given that Arabic geminates are spelled with a diacritical sign and a single consonant. Arabic speakers tend to convert the spelling of Arabic geminates into a single consonant in English because Arabic geminates are spelled with a singular consonant and a diacritical mark that is not frequently shown in the written version used by Arab seniors. Additionally, when transliterating Arabic names into English that are sounded with a single consonant phoneme, they overgeneralized the use of double consonants. There are suggestions for enhancing native Arabic speakers' ability to transliterate their names on social networking sites.

Proper name lexicons that are bilingual are essential for cross-language information retrieval and machine translation. Typically, word alignment techniques are applied to automatically create bilingual lexicons from parallel corpora. Proper name alignment becomes especially challenging when there is a difference in written script between the source and target languages of the parallel corpus. In this study Semmar and Saadane (2013) provided an automatic transliteration method for proper names from Arabic to Latin script, as well as an alignment tool for single and compound words from parallel English-Arabic texts. They paid special attention to how transliteration can enhance the word alignment tool's functionality. Two approaches have been used to assess the word alignment tool that integrates proper name transliteration from Arabic to Latin script: An assessment of the alignment quality by hand as well as an analysis using the open-source statistical machine translation program Moses to determine how this alignment affected the quality of the translation. Test results indicate that adding proper name transliteration to the alignment procedure increases the translation BLEU score from 20.15% to 20.63% and the F-measure of word alignment from 72% to 81%.

Al-Jarf (2022 b) sought to determine the distribution of English alternative transliteration of Arabic names, the kinds of variations transliterations produced by Arabic speakers, the methods they employ when transliterating their names to English, and the sources of variations in the English transliteration of the same Arabic name. It also sought to explore variant transliterations of the same Arabic names in English by Arabic native speakers on Facebook. The author's Facebook friends provided a sample of 112 names, in all 332 variation transliterations and 1139 occurrences (repeats). It was discovered that 26% of Arabic names have three different English transliterations, while 59% of Arabic names have two. Names with the highest number of variant transliterations are الجرف (35), محمد (7); يوسف, شريف محمود, (6); نور نورة شيماء (5). Variants with the highest occurrences are الجرف (154), محمد (153); احمد (90); Ali (67); محمود (53); هناء & السيد(91). .

Arabic lexical ambiguity presents several difficulties for machine translation systems, which has an impact on the precision and quality of their translations. Aldawsari (2023) investigated how SYSTRAN and Google Translate, two well-known MT systems, handle homonyms, heteronyms, and polysemes—three challenging linguistic aspects of the Arabic language. Sentences with these properties in a variety of contexts and domains were included in a test suite. Four independent assessors used a four-point rating system to assess the intelligibility and correctness of the translations generated by the two MT systems. The three selected linguistic features proved difficult for both MT systems to master, as evidenced by their average scores being below 40%. In nearly every statement in the test suite, Google Translate performed better than SYSTRAN. Intelligibility was a higher score for both systems than accuracy. Due to the Arabic discretization's distinctive architecture, which translation algorithms are yet unable to recognize, heteronyms proved to be the most difficult for both MT systems. This work advances the subject of machine translation by offering a thorough examination of Arabic lexical ambiguity and how it affects the quality of machine translation systems. It also makes recommendations for potential enhancements to these systems for translating Arabic into English.

The goal of Almahameed et al. (2017) was to investigate how various Arabic terms and phrases are translated into English for use in travel fiction publications. This was accomplished by looking at Arabic words and phrases that have been transliterated and contrasting the various transliterations of identical words and phrases made by various authors. It investigates how travel writers create plural nouns, how to utilize the definite article (Al) (ال), and how to deal with Al-tashdid, which is the repetition of a letter. The findings are that transliteration is a common tool used by writers of travel literature for three main reasons: first, it can be used when there isn't always a direct correlation between words; second, it can help prevent meaning loss; and third, it can facilitate cross-cultural transculturation. Academics who study the transliteration of Arabic into English, particularly in English travel literature, will find this essay's contributions to be of interest. As far as scholars are aware, this essay is a trailblazing investigation of the transliteration of Arabic into English in English travel literature.

The transliteration technique used to represent Arabic characters in the book is introduced in the chapter authored by Habash et al. (2007). The system is a comprehensive, readable, one-to-one transcription of the Arabic alphabet that complies with Arabic computer encoding standards. Using this transliteration scheme, we offer pronunciation instructions for Arabic and talk about the peculiarities of Arabic spelling.

Previous research in Arabic translation and transliteration includes Jabak (2023) on Omani students' English news translation errors, Al-Jarf (2022a) on Arabic name transliteration variations, Semmar and Saadane (2013) enhancing bilingual lexicon creation, and

Aldawsari (2023) exploring MT challenges with Arabic lexical ambiguity.

3. Methods

The study employed a mixed methods approach, gathering transliteration data from Google Translate, a widely used free online translation service in Saudi Arabia. It was conducted at a public institution in the central region of Saudi Arabia during the first semester of the academic year 2023-2024.

For qualitative insights, voluntary interviews were conducted with six specialist translators who were professionals in the field. All participants were Saudi citizens with over 10 years of professional experience. Each interview was recorded with the participant's consent and transcribed to identify trends in transliteration practices within translation.

Regarding the quantitative aspect, the study focused on transliteration challenges in Arabic to English translation efforts. A meticulous examination of a diverse parallel corpus of Arabic texts with their English transliterations was conducted to highlight various aspects of transliteration ambiguity, such as phonetic variations, context-dependent mappings, and lexical homophony. Both manual and machine-based analysis approaches were employed to gather data, with a focus on identifying instances of transliteration ambiguity. Real-world transliteration examples from students' translation outputs were utilized to illustrate the practical implications of ambiguity in Arabic-English transliteration.

4. Results and Discussion

RQ1: What are the primary sources of ambiguity in the transliteration of Arabic words/phrases, into English?

The following leading themes were identified from the transcripts of the interviews to answer the first research question:

1. The practice in non-standardized or informal communication is in replacing graphemes rather than complete word/ lexical units during translation. This has come to be known as Arabizi or Arabish from Arabic + English (*Angrezi*). This is especially useful when the materials being translated target the second generation bi or multilingual Arabic speakers and their likes, a communication that assumes reasonable English proficiency in the target population, is anchored in English-based technology, and is dynamic to a large extent. Graphemes such as ح (Haa), which is a stronger "Hh" sound than the ه (haa), are typically replaced by similar looking characters on the regular English keyboard/ keypad such as '7', and '3' for ع ('ayn).
2. In standardized or formal translation, where the translator uses the translation apps or tools, the most drastic variation is found in some place names: frequently occurring place names such as Mecca find variations such as Makkah and Mekkah,
3. In standardized or formal translation based on online tools, stock Arabic phrases such as the greeting, Assalam Alaikum find different variations such as assalamuʿalaykum, *assalamu alaikum*, as-Salaam-Alaikum, and the most common Arab name has many variations such as Muhammad, Mohammed, Mohamed, Mahomet.
4. As a practice in translation, transliteration is only occasionally resorted to by Saudi translators during manual processing of documents, yet it poses a great problem when English is the target language. One of the specialists added that despite very careful revision of translated documents, sometimes lack of graphemic equivalence in the target language forces such profound changes in proper nouns that the final product is totally unrecognizable.

   Thus, to answer the first research question, the primary sources of ambiguity in transliteration from Arabic to English is presented by non-standardization of practices. The rapid pace of globalization and technological tools such as chatting have largely contributed to this flexibility in transliteration. It is both a boon and a bane: Former as it creates opportunities for communication especially in the new paradigm of cryptic messaging. At the same time, it is a bane as it has created two approaches even within the realm of occasional transliteration: One, the informal approach where symbols such as numerics are used to denote graphemes, a truly unique mixing of 'scripts' between the Arabic-English language pair. Two, machine translation led to greater incidence of ambiguity in Arabic to English transliteration as at grapheme, phonology, and morphological levels, equivalence was challenged. These findings align with some previous literature. For example, Al-Salman (2008) assigned the problems in transliteration to the uniqueness of each of these languages and the sounds that they are composed of. Similarly, Al-Sughaiyer and Al-Kharashi (2004) reported that the morphological complexities of Arabic, where word formation and the addition of affixes add to the ambiguity during transliteration.

RQ2: How do phonetic variations in Arabic dialects contribute to transliteration ambiguity into English?

It is found that the divergence among these Arabic varieties is influenced by the historical and present impact of other languages spoken in their respective regions**.** Closely related forms of Arabic generally share mutual intelligibility, but more distant variants often lack this quality. The differences are especially pronounced in varieties to the west of Egypt. South Arabian, meanwhile, in present day Iraq, the languages Syriac Aramaic, Akkadian, Babylonian, and Sumerian have played a role, and Persian has had an impact in the Middle East. Most translation tools work by figuring out the most relevant translation of content in a broad Egyptian Arabic speakers, for instance, often express challenges in comprehending North African Arabic speakers. On the other hand, the ability of North African Arabic speakers to understand their peers from different regions is primarily attributed to the widespread influence of Egyptian Standard Arabic, and to a lesser degree, the prevalence of media from the Levantine region.

The divergence among these Arabic varieties is influenced by the historical and present impact of other languages spoken in their

respective regions. In Egypt, for example, influences come from languages like Coptic, Greek, and English. North Africa and the Levant, on the other hand, have been shaped by the languages French, Ottoman Turkish, Italian, Spanish, Berber, Punic, or Phoenician. In Yemen, linguistic influences include Himyaritic, Modern South Arabian, and Old context and then, rearranging and adjusting them to replicate human speech with correct grammar. Dialectal variations in Arabic, therefore, are ignored for the Modern Standard Arabic version of the language, leaving some scope for transliteration ambiguity especially when one is looking for translation from direct speech. These finding agreed with Prior et al. (2011) who reported that ambiguity in some syntactical patterns within the same dialect has always been problematic for automated translation systems as well as translators who are humans. Furthermore, Al-Muhtaseb and Mellish (1998) stated that current computational models for Arabic-English transliteration often lack the nuance and adaptability required to handle real-world texts, especially when juggling between formal Modern Standard Arabic and regional dialects. This underscores the need for advanced models that are both robust and adaptable, equipped with a comprehensive understanding of ambiguity markers and their implications

RQ3: What role does lexical homophony play in introducing transliteration uncertainties?

To check out how machine tools dealt with Arabic homophones in translation, the following words were translated on Google Translate, the most popular and frequently used translation tool:

عَيْن (spring/ eye); عَامِل (factor/ worker); قَرْن (century/ horn); تَارِيخ (date/ history); خِيَار (cucumber/ choice); تَذْكِرة (reminder/ ticket); هَاوٍ (falling/ amateur); تالٍ (next/ reader); أَهْل (relatives/ qualified); حَيّ (neighbourhood/ alive); مَكْتَب (desk/ office); عَرْض (show/ offer); ظَرْف (envelope/ circumstance); صَاحِب (owner/ friend); ذَهَب (gold/ he went); سَهْل (plain land/ easy); دَلِيل (guide/ proof); أَدَب (good manners/ literature); سَائِل (liquid/ questioner); طَرْد (package/ expulsion).

Without exception, the tool only produced one of the two possible meanings for the Arabic homophones, thus establishing the fact that homophony can contribute to transliteration uncertainties when translators rely on online tools. However, during the interviews when the six translators were questioned on the problems that Arabic homophones posed, they unanimously agreed that manual transliteration was almost completely unlikely to cause such ambiguity. These reported findings are confirmed with previous studies (Gugliotta, 2022; Jabak; 2023; Zemni et al., 2024). To begin Gugliotta's (2022) findings which reported that lexical ambiguity of Arabic in NLP presents extra challenges. When two or more Arabic words translate to a single English word, this is frequently the case. Zemni et al. (2024). Jabak (2023) found that most of the mistakes made by the students were in the areas of lexical and grammatical errors, acronyms, cities, and correct adverbs. The findings of this study are likely to offer valuable insights to professionals involved in cross-cultural communication, localization, and information retrieval, as well as to developers of transliteration tools and systems. In the long term, these findings can help deeply comprehend ambiguity markers in the Arabic-English language pair transliteration process and make users more aware of the same, fostering effective cross-cultural communication and knowledge dissemination.

5. Conclusions

The practice of grapheme replacement in informal communication, known as Arabizi or Arabish, reflects an adaptive approach to cater to the preferences and proficiency levels of second-generation bilingual or multilingual Arabic speakers. Informal communication strategies demonstrate flexibility and adaptability to the linguistic dynamics and preferences of the target audience. The infrequent use of transliteration by Saudi translators in formal translation processes, coupled with the challenges it presents, underscores the complexity of maintaining accurate representations, especially when English is the target language. Transliteration requires careful consideration, and the challenges faced in maintaining graphemic equivalence highlight the potential for changes in proper nouns during formal translation. Standardized or formal translations, whether through apps or online tools, exhibit variations in place names and stock Arabic phrases. This variability suggests that maintaining consistency in transliteration and standard expressions can be challenging. Standardized translation processes may benefit from additional measures to ensure uniformity, particularly in the representation of frequently occurring terms and phrases. The dichotomy between informal grapheme replacement and formal transliteration underscores the need for translators to strike a balance between catering to dynamic, English-centric audiences and ensuring accuracy in formal, standardized contexts. Translators must navigate between linguistic adaptability and the preservation of linguistic and cultural integrity, especially in diverse translation scenarios. These findings highlight the dynamic nature of Arabic-English translation, where informal and formal practices each have their nuances and challenges. The findings emphasize the importance of considering the target audience, linguistic adaptability, and maintaining consistency in formal translation processes.

The linguistic landscape of Arabic is diverse, with variations influenced by historical, geographical, and cultural factors. Understanding these variations is crucial for effective communication within the Arabic-speaking world. The challenges in translation tools highlight the importance of recognizing and preserving dialectal nuances in Arabic. Translators need to be mindful of the specific context and linguistic variations to ensure accurate and culturally relevant translations. While Modern Standard Arabic serves as a unifying standard, acknowledging and incorporating dialectal differences is essential for authentic and contextually relevant communication. Balancing standardization with the preservation of linguistic authenticity is crucial in translation processes. The study underscores the limitations of popular online translation tools, such as Google Translate, in dealing with homophones in Arabic. This limitation may impact the accuracy of transliterations, particularly in languages with rich linguistic nuances. The unanimous agreement among specialist translators suggests that human expertise plays a crucial role in mitigating transliteration ambiguities, particularly when dealing with homophones. Manual transliteration is viewed as a more reliable approach, emphasizing the importance of human involvement in nuanced language translation.

The results emphasize the need for caution when relying solely on machine translation tools, especially for languages with homophones. Human translators bring a level of understanding and context that automated tools may lack. The study highlights the transliteration challenges posed by Arabic homophones when using online tools, emphasizing the reliability of human translators in handling such linguistic nuances with greater accuracy and clarity. This study is a unique contribution in the field of translation studies as it sheds new light on ambiguity markers in th sub-field of transliteration between the Arabic-English language pair. holds paramount importance in advancing the field of transliteration studies. The study is also of primary use to professionals who use transliteration for communication purposes such as those in the tourism industry, A thorough understanding of ambiguity markers in Arabic-English transliteration will likely equip them in cross-cultural communication, localization, and information retrieval with the knowledge to foster more accurate exchanges. Additionally, for developers creating transliteration tools, the findings will provide crucial insights for designing systems that are adaptable, nuanced, and context-aware, thus promoting effective cross-cultural communication.

Transliteration is a fast evolving and highly dynamic field of study given the fact that technological tools and the tendency for cryptic communication are both contributing to the birth of new communication possibilities in languages across the globe. This study had a limited scope in the sense that it was based on the Arabic to English transliteration challenges and findings. However, since different languages are likely to act differently especially in online translation, there is a need for replications across language pairs to make the results generalizable.

Acknowledgments

The authors would like to thank The Arab Observatory for Translation (an affiliate of ALECSO), which is supported by the Literature, Publishing & Translation Commission in Saudi Arabia.

Authors contributions

Dr. Majed Abdullah Alharbi was responsible for research design, data collection, research methodology, conclusion and revision of the article. Dr. Mohammad Shariq finalized literature review, data analysis, recommendations and references. Both authors reviewed the final version of the article and made necessary adjustments time to time.

Funding

This research received grant no. (143/2023) from the Arab Observatory for Translation (an affiliate of ALECSO), which is supported by the Literature, Publishing & Translation Commission in Saudi Arabia.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Aldawsari, H. A. (2023). Comparing the Performance of Google Translate and SYSTRAN on Arabic Lexical Ambiguity. *Arab World English Journal for Translation and Literary Studies, 7,* 19-34. http://doi.org/10.24093/awejtls/vol7no3.2

Al-Jarf, R. E. (2022a). Gemination errors in Arabic-English transliteration of personal names on Facebook. *International Journal of*

Linguistics Studies, *2*(2), 163-170. http://doi.org/10.32996/ijls.2022.2.2.18

Al-Jarf, R. E. (2022b). Variant transliterations of the same arabicArabic personal names on Facebook. *International Journal of English Language Studies*, *4*(4), 79-90. http://doi.org/10.32996/ijels.2022.4.4.11

Almahameed, N. A., Abbadi, R. M., & Almahameed, A. A. (2017). Between languages and cultures: Arabic into English transliteration in English travel literature. *International Journal of Applied Linguistics and English Literature*, *6*(6), 235-245. https://doi.org/10.7575/aiac.ijalel.v.6n.6p.235

Al-Muhtaseb, H., & Mellish, C. (1998, April). Some differences between Arabic and English: A step towards an Arabic upper model. *In 6th International Conference on Multilingual Computing*.

Al-Salman, A. (2008). Phonetic variations in Arabic-English transliteration. *Linguistic Journal*, *4*(2), 45-58.

Al-Sughaiyer, I. A., & Al-Kharashi, I. A. (2004). Arabic morphological analysis techniques: A comprehensive survey. *Journal of the American society for information science and technology*, *55*(3), 189-213. https://doi.org/10.1002/asi.10368

Askari, S. (2019). *Stereotyping and translation in Arabic and English news texts with reference to Islamophobia and the Arab-Israeli conflict* (Unpublished PhD dissertation). University of Leeds.

Deheuvels, L. W., Michalak-Pikulska, B., & Starkey, P. (Eds.). (2006). *Intertextuality in modern Arabic literature since 1967* (Vol. 2). Durham Modern Languages.

El Kholy, A., & Habash, N. (2012). Orthographic and morphological processing for English–Arabic statistical machine translation. *Machine Translation*, *26*, 25-45. https://doi.org/10.1007/s10590-011-9110-0

Gugliotta, E. (2022). *Tunisian Arabizi: Linguistic analyses and corpus building using natural language processing*. (Unpublished PhD dissertation). Universit éGrenoble-Alpes; Sapienza Universita di Roma.

Habash, N., Soudi, A., & Buckwalter, T. (2007). On Arabic transliteration. In *Arabic computational morphology: Knowledge-based and empirical methods* (pp. 15-22). Publisher: Springer. http://doi.org/10.1007/978-1-4020-6046-52

Halimah, A. M. (2016). Translating Arabic proper names: A foreignising approach. *International Journal of English Language and Linguistics Research*, *4*(2), 1-16.

Hassan, H. (2019). Investigating the challenges of translating Arabic collocations into English with reference to the Quran (Unpublished PhD dissertation). Liverpool John Moores University.

Issa, I. (2018). *The morphological complexity of L1 Arabic-speaking children* (Unpublished PhD). Rhodes University.

Jabak, Omar. (2023). Analysis of errors made by a sample of Omani students in translating englishEnglish news headlines into Arabic. *Canadian Journal of Language and Literature Studies*, *3*(6), 20-34. http://doi.org/10.53103/cjlls.v3i6.131

Masmoudi, A., Khmekhem, M. E., Khrouf, M., & Belguith, L. H. (2019). Transliteration of Arabizi into Arabic script for Tunisian dialect. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, *19*(2), 1-21. https://doi.org/10.1145/3364319

Prior, A., Wintner, S., MacWhinney, B., & Lavie, A. (2011). Translation ambiguity in and out of context. *Applied Psycholinguistics*, *32*(1), 93-111. https://doi.org/10.1017/S0142716410000305

Rees, G. P. (2018). *A phraseological multi-discipline approach to vocabulary selection for English for academic purposes* (Unpublished PhD dissertation). Universitat Pompeu Fabra.

Regmi, K., Naidoo, J., & Pilkington, P. (2010). Understanding the processes of translation and transliteration in qualitative research. *International Journal of Qualitative Methods*, *9*(1), 16-26. https://doi.org/10.1177/160940691000900103

Semmar, N., & Saadane, H. (2013, October). Using transliteration of proper names from Arabic to Latin script to improve English-Arabic word alignment. *In Proceedings of the Sixth International Joint Conference on Natural Language Processing* (pp. 1022-1026).

Zemni, B., Zitouni, M., Bouhadiba, F., & Almutairi, M. (2024). On Ambiguity in the Arabic Language: Scrutinizing Translation Issues through Machine Translation from English and French into Arabic. *Journal of Intercultural Communication*, *24*(1), 203-212. https://doi.org/10.36923/jicc.v24i1.171

Томін, А. В. (2023). *Features of the translation of phraseological units in contemporary fictional discourse based (on modern American literature).* Kyiv National Linguistic University.