Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Arabic Transliteration of Borrowed English Nouns with /g/ by Artificial Intelligence (AI)

Prof. Reima Al-Jarf
King Saud University, Riyadh, Saudi Arabia
Corresponding Author: Prof. Reima Al-Jarf, **E-mail**: reima.al.jarf@gmail.com

| ABSTRACT

This study sought to explore the transliteration of English nouns containing /g/ by Microsoft Copilot (MC) and Google Translate (GT), and find out which Arabic equivalents are transliterated with ك ,ق ,ج ,غ, which nouns have two or more transliterations, causes of transliteration variations in Arabic, and comparison of AI and human transliteration accuracy. A sample of 140 English nouns containing the letter g pronounced /g/ was collected and transliterated by MC and GT. It was found that MC and GT rendered almost identical Arabic transliterations of English names with /g/ in the sample. Arabic transliterations by MC and GT were almost identical to those of human transliterations. In 61% of the English nouns, MC and GT used غ /ɣ/ to represent /g/, of which غ is pronounced /ɣ/ (17%) as Singapore سنغافورة, Chicago شيكاغو, Ghana غانا, Congoالكونغو , and غ is pronounced /g/ in 44% (Angola أنغولا, Gambia غامبيا , Gluten غلوتين). In 37%, the Arabic equivalents are transcribed with ج, where ج is pronounced /g/ in 27% (Copenhagen كوبنهاجن, Craig كريج, Godiva جوديفا, Gordon جوردون, Greenland جرينلاند) and ج is pronounced /dʒ/ in 10% (Colgate كولجيت, Cigarettes السجائر, Galaxy جالكسي, Gallon جالون). Use of غ is more common in the Levant whereas ج is more common in Egyptian Arabic. MC and GT gave a single Arabic transliteration to each English noun, whereas human transliterators yielded two or more transliterations to 16% of the sample (Google قوقل جوجل جوجل كوكل غوغل گوگل چوچل; Group غروب جروب، قروب، گرو; English الإنجليزية الأنقليزية الانكليزية; Hashtag هاشتاج هاشتاق هاشتاغ). Speakers of Saudi Arabic use ق (Goody قودي, Oregano اورقانو, Golf قولف, Grand Store قراند ستور, grammar قرامر, Reading ريدينق). Only one case was transliterated with ق by MC (hashtag هاشتاق), and Guilin قويلين by GT. Transliteration variations in Arabic are caused by an interplay of factors as regional and dialectal variations, lack of knowledge, widely recognized, old, established transliterations of names, especially in the geopolitical domain (Chicago شيكاغو), lack of standardization and policy in Arab countries, and contextual and pragmatic factors. The article concludes with some recommendations for the accurate Arabic transliteration of foreign names.

| KEYWORDS

Letter/sound g, proper nouns, transliteration of names, English-Arabic transliteration, Artificial Intelligence, AI transliteration, Human transliteration

| ARTICLE INFORMATION

ACCEPTED: 01 August 2025 **PUBLISHED:** 02 September 2025 **DOI:** 10.32996/jcsts.2025.7.9.29

1. Introduction

English has consonant phonemes and letters that are non-existent in Arabic as /p/, /v/, ch /tʃ/ and /g/. In the transliteration of English words borrowed in Arabic such as nouns referring to personal, city, country, brand, shop names and others, these phonemes/letters are usually substituted by certain Arabic letters that represent the closest match. In some cases, the same English phoneme is represented by several Arabic letters depending on the Arabic vernacular. For example, in the transliteration of foreign words containing the sound /g/, the letter غ /ɣ/ is preferred in the Levant (nowadays), foreign publications and Arabic TV Channels as Aljazeera, al-Mayadeen, Al-Arabi, Deutsche Welle, and Alhurra to represent /g/ as in Hong Kong > هونغ كونغ , Ghana > غانا, August > أغسطس, and Garcia> غارسيا, Agatha Christi أغاثا كرستي. In some nouns, it is often pronounced /g/ (Bangladesh بنغلاديش, Guatemala > غواتيمالا), and in others, it is pronounced /ɣ/ (Chicago > شيكاغو, Ghana > غانا. In Egypt, where ج is pronounced /g/ in all situations, ج is used to transcribe foreign /g/ in all contexts (gallon>جالون, glucose > جلوكوزو ,

Gateau > جاتوه, Gloria Jean > جلوريا جين) and so on. In coastal Yemen, southwestern and eastern Oman, in Hejazi and Najdi Arabic, it is ق /q/ as in Goody > قودي, Grammar > قرامر. In Iraq, گ is more used in *England* > إنكلترا. In Tunisia and Algeria, ڤ or *qāf* ق is used. In Morocco, ڴ or ك is used. In Lebanon, the letter چ is often used to create the phoneme /g/ in foreign loanwords, such as in چوچل (Google)[1]. Words as "*group*", which is used in spoken Arabic has four transliterations غروب گروب قروب، جروب , and *Google* has seven transliterations (كوكل غوغل گوگل چوچل جوجل قوقل جوجل) (Al-Jarf, (2015), Al-Jarf, 2003; Al-Jarf, 1995; Al-Jarf, 1990; Al-Jurf, 2002; Al-Jurf, 1994).

The transliteration of foreign/English words containing /g/ to Arabic has been of interest to researchers for decades. For example, Aziz (1983) offered one of the earliest systematic treatments of English-to-Arabic transliteration. He analyzed phonological mismatches, script limitations, and the role of diacritics. His study proposed transliteration rules for consonants and vowels which included /g/ that is often rendered as ج (in Egyptian Arabic) or ك، ق، غ depending on the dialect and domain *Gordon* > غوردون (in formal contexts); /v/ > ف, since Arabic lacks /v/ as in *Victor* > فيكتور; /p/ > ب, due to absence of /p/ as in *Paul* > بول. Short vowels in English are often omitted in Arabic script unless it is necessary to add vowels for clarity (*Helen* > هيلن not هيلين; *Kevin* > كيفن, not كيڤن). Aziz warned against over-literal spelling that misleads pronunciation. He stressed the importance of consistent transliterations across texts and contexts. *Michael* should always be مايكل, not ميكائيل or ميشيل) unless context demands religious or cultural adaptations. Aziz critiqued the lack of unified transliteration norms across Arab countries, calling for institutional standardization.

The inconsistency and inaccuracy in English-Arabic transliteration was critiqued by Odisho (1992), especially for sounds like /g/, /v/, and /p/. He compared regional practices (e.g., Egyptian use of ج, Iraqi use of گ,) and called for phonetic precision using Arabic script's full potential. He showed how transliteration functions as a semiotic and communicative act, reshaping meaning as English lexical items are encoded into Arabic script. Semiotic distortions arise when phonemes like /p/ and /v/, absent in Arabic, are substituted with /b/ and /f/, leading to symbolic mismatches. For instance, "computer" becomes كمبيوتر, and "club" is rendered as كلب, risking confusion with the Arabic word for "*dog*." Odisho highlighted how such substitutions compromise indexical clarity, especially in brand names and technical terms, where the original referent may be obscured. Inconsistent transliteration such as "garage" as كراج or "bank" as بنك, can hinder intelligibility and semantic precision. Odisho emphasized that transliteration is not merely phonetic mapping but is also a culturally loaded process that affects how foreign concepts are integrated, perceived, and circulated within Arabic discourse.

In transliterating Omani place names into English for use on maps and signages, Muscat (2011) examined inconsistencies, due to phonological mismatches, lack of diacritics, and dialectal variation as in the following examples: الخوض (*Al Khaud*, *Al Khoudh*, *Al Khawdh)*; قاسم (*Qasim*, *Gasim*, *Jasim*, *Kasim)* where the /q/ sound is variably rendered as /g/, /j/, or /k/; السيب (*Seeb*, *Sib*, *As Seeb)* with variations in representing the definite article "ال"; صلالة (*Salalah*, *Salala)* with vowel length and gemination inconsistently represented; and نزوى (*Nizwa*, *Nazwa)*. These examples show variations in transliteration across ministries, maps, and signage, reflecting lack of standardization, dialectal influence and simplification of names on tourist maps for readability, at the expense of linguistic accuracy. He proposed a simplified transliteration system that blends onomastics, phonology, and national identity planning. The influx of non-Arabic speakers pushes for simplified, readable English forms. Transliteration should take into consideration Arabic sounds like /ʕ/, /ɣ/, and emphatic sounds that have no English equivalents. Since, place names carry historical and symbolic weight, transliteration choices can dilute or distort that. The author recommended standardizing transliteration to preserve the Omani cultural heritage and to improve tourists' decoding ability. He also advocated the Unified Arabic Transliteration System (2007 Beirut System) and the guidelines set by the United Nations Group of Experts on Geographical Names (UNGEGN) for the transliteration of geographical names.

According to Al-Omar, 2013, English-Arabic transliteration of proper nouns is a difficult process due to the idiosyncratic systems of the phonemic inventories, syllabic structure constraints, lexical stress in both languages, and dialectal variations. For example, the English name "*Maggie*" is transliterated into two different forms in Arabic: One with letter غ, resulting in "ماغي", the other with letter ج, creating ماجي. "*Bergman*" is transliterated with ج (برجمان), *Agatha* is transliterated as "أغاثا. The English consonant /p/ in *Peter* is often transliterated with ب /b/; /v/ in *David* is often transliterated using the Arabic consonant ف /f/; the English palato-alveolar affricate /ʧ/ ('ch' sound) in *Churchill* is substituted with تش (تشرتشل). The author also pointed out the historical practice of substituting sounds that do not exist in standard Arabic with similar-sounding Arabic letters. These substitutions often lead to variations in transliteration over time. Al-Omar emphasized that transliteration must be approached with linguistic rigor and cultural awareness, advocating for standardized systems that preserve phonetic integrity while facilitating intelligibility and mutual respect between Arabic and English-speaking communities.

[1] Ghayn - Wikipedia

Drawing on examples from Arabic, Jāwī, Urdu, Korean, and Russian, Gu (2025) revealed how transliteration localizes English while preserving its global identity. English words are re-scripted phonetically, semantically, and symbolically, often passing as native terms (بيتزا *pizza,* برقر *burger, Orange* أورانج). These examples also demonstrate how multiscriptal English functions as a dynamic sociolinguistic resource, shaped by local orthographic norms, cultural filters, and communicative priorities. They reveal how English is "glocalized," acquiring new symbolic and functional roles in multilingual societies. By inhabiting local scripts, English gains new symbolic and communicative functions, reflecting broader processes of globalization and cultural hybridity.

The transliteration of English loanwords into Arabic in the media and education was investigated by Haider et al. (2025) who examined how borrowed words (loanwords) from English are integrated into Arabic through phonetic transliteration (prioritizing pronunciation) and orthographic transliteration (prioritizing spelling conventions). They showed that transliteration inconsistencies arise from lack of standardization, dialectal variation, and semantic domain (*club* كلوب*, computer* كمبيوتر*, Tweeter* تويتر*, Google* جوجل*, video* فيديو *, email* إيميل *, software* سوفتوير*, burger*برقر*, pizza* بيتزا) . The authors proposed a hybrid model combining orthographic familiarity and phonetic accuracy and highlighted the role of institutional norms and audience expectations. The study offers practical guidelines for translators for handling borrowed terms.

A group of studies by the author on the transliteration of personal names on social media and shop, brand, product, and hotel names focused on absence of vowels in English spelling of Arabic personal names as *Ahmad* may appear as "*Ahamd*", "*Ahmd*", or "*Amd* (Al-Jarf, 2023a); English Spelling of Arabic compound personal names like "*Abdul Rahman*" appear in multiple variants: *Abdulrahman, Abd Alrahman, Abdelrahman, Abdalrahman* and inconsistent attachment of prefixes (*Abd, Abu*) and suffixes (*ddine, din, dine*) (Al-Jarf, 2023b); spelling of glottal stop and pharyngeal fricative in Arabic names containing ع /ʕ/ and ء /ʔ/ (Al-Jarf, 2023c); transliteration of Arabic personal names with the definite article {al-} as *Aljarf, Al Jarf, Al-Jarf, Elgarf, Ljarf* (Al-Jarf, 2022b); gemination errors in Arabic-English transliteration of personal names like *Nassar, Algammal,* and "*Alqattan* are often misspelled as *Nasar, Algamal, or Alqatan*. Conversely, non-geminated names like "*Ahmad*" are incorrectly doubled as "*Ahmmed* (Al-Jarf, 2022c); variant transliterations of the same Arabic personal names like "*Mohammad*", "*Muhammad*", "*Mohamed*", and "*Mohamad*" show high variability (Al-Jarf, 2022d); faulty transliteration of some hotel names in Saudi Arabia (Al-Jarf, 2021b); similar, transliteration anomalies in compound shop names like "*Operation Falafel*" (فلافل أوبريشن) and "*Rude Shake*" (رود شيك) (Al-Jarf, 2023e); deviant Arabic transliterations of foreign shop names in Saudi Arabia as *Burger King* rendered بيرغر & برجر; كينج & كنج (Al-Jarf, 2022a). Errors in transliterating personal, shop, brands, product, and hotel names in these studies are influenced by dialectal pronunciation and habits, lack of awareness of English vowel representation, lack of standardization, confusion over capitalization and spacing, and transfer from Arabic script, where gemination is marked by a diacritic not shown in casual writing.

Regarding the development and use of AI transliteration models, AbdulJaleel & Larkey (2003) introduced a statistical model for transliterating English proper nouns into Arabic to improve cross-language information retrieval. The authors used a Hidden Markov Model (HMM) trained on transliteration pairs to predict Arabic equivalents. They addressed challenges like short vowel omission and multiple valid transliterations. Their system outperforms dictionary-based methods in precision and recall. The study emphasizes transliteration as a bridge for query expansion in bilingual search engines. It remains influential in Arabic Natural Language Processing (NLP) and Information Retrieval (IR) systems. Likewise, Fattah & Ren (2008) developed algorithms to extract transliteration pairs for proper nouns. They addressed challenges like short vowel omission and low-frequency names. Their system uses similarity metrics and pattern matching to align English and Romanised Arabic names. The study improves bilingual lexicon building for machine translation and information retrieval. It's a technical contribution to computational linguistics and transliteration automation. In addition, absence of Arabic letters for sounds like /g/, /p/, and /v/ deters accurate transliteration of global terms in smart city communication systems (Asar & Abdullah, 2025).

The above literature review shows lack of studies that examine the transliteration of personal, city, country, product, and shop names and other nouns by Artificial Intelligence (AI) tools/assistance in comparison with human transliterators. Therefore, this study aims to explore the following: the transliteration of personal, city, country, product, brand, and shop names containing /g/ in initial, medial and final positions by Microsoft Copilot (MC) and Google Translate (GT), which Arabic equivalents are transcribed with غ, ج, ق, or ك, which ones have two or more Arabic equivalents, causes of variations in the transcription of English names with /g/ to Arabic غ، ج، ق، ك, accuracy of AI transliteration compared with that of human transliterators, and whether translation students and social media users can use AI tools/assistance as MC, GT, Deepseek, Gemini, META AI, and others in transliterating personal, city, country, product, brand and shop names from English to Arabic.

This study is significant because, in this day and age, Arab social media users make a lot of spelling and transliteration mistakes on social media, and they completely ignore Standard Arabic spelling rules. They spell words the way they pronounce them in their local dialects. They misspell personal, city, and country names that have had a fixed spelling for a very long time (Al-Jarf, 2023d). This study will also raise social media users and student translators' awareness of the standard spelling used in the English-Arabic transliteration of personal, city, country, brand, product, and country names. Communicatively, English-Arabic

transliteration of names targets people who know the English names aurally but do not know how they are transcribed in Arabic, especially in the case of names containing the letter g which is pronounced /g/.

2. Data Collection and Analysis

A sample of 140 English nouns containing the letter g pronounced /g/ in initial, medial or final positions was collected. The nouns refer to personal, city, country, brand, shop, and hotel names and others. The nouns and their commonly used Arabic transliterations by humans were collected from Arab mainstream media, social media, print materials, Arabic maps, searching Google for specific noun categories (city names, countries, shop names with letter g), in addition to the author's pictures and lists of shop, hotel, brand and product names and linguistic landscapes in her previous studies. The English names in the samples were transliterated by Microsoft Copilot (MC) and Google Translate (GT). Five colleagues specialized in Linguistics, Arabic language, geography and business verified the Arabic transcriptions rendered by MC and GT, and judged their accuracy and compatibility with human transliterations. Discrepancies in judgments were solved by discussion. Percentages of English names in which g was transliterated as غ & ج and the percentages of names in which غ is pronounced /g/ & /ɣ/ and the percentages of names in which ج is pronounced /d͡ʒ/ & /g/ were calculated. Results of the analysis are reported quantitatively and qualitatively.

3. Results

Data analysis showed that both MC and GT rendered almost identical Arabic transliterations to all English names with /g/ in the sample. In addition, the Arabic transliteration by MC and GT was almost identical to that of human transliterators.

Additionally, it was found that in 61% of the English nouns in the sample, غ /ɣ/ was used to represent /g/. These have two subsets. In the first subset (17%), غ is pronounced /ɣ/ by Arabic speakers regardless of their dialect as in *Afghanistan*افغانستان, *Afghan* افغان, *Chicago* شيكاغو, *Congo* الكونغو, *Erdogan* اردوغان, *Gaziantep* غازي عنتاب, *Geographic* جغرافي, *Ghanan* غانا, *Gorilla* غوريلا, *Gram* غرام, *Guinea* غينيا, *Madagascar* مدغشقر, *Magnesium* مغنيسيوم, *Moghul* مغول, *Mongolia* منغوليا, *Portugal* البرتغال, *Pythagoras* فيثاغورس, *Senegal* السنغال, *Singapore* سنغافورة, *Uganda* اوغندا, *Agatha Cristi* أغاثا كريستي, *gas* غاز. All of these examples are Arabizations of the English source name and have been pronounced with /ɣ/ for a very long time.

In the second subset (44%), غ is pronounced /g/ by Arabic speakers as in *Angola* أنغولا, *Bangladesh* بنغلاديش, *Bengal* البنغال, *Birmingham* برمنغهام, *Bogota* بوغوتا, *Douglas* دوغلاس, *Gaborone* غابورون, *Gabriel/Gabriela* غابرييل غابرييلا, *Gambia* غامبيا, *Garcia* غارسيا, *Gary* غاري, *Gdynia* غدينيا, *ginseng* الجينسنغ, *Glasgow* غلاسكو, *Gluten* غلوتين, *Gold Coast* غولد كوست, *Golda Mair* غولدا مائير, *Gracia* غراسيا, *Graffiti* غرافيتي, *Grand Rapids* غراند رابيدز, *Graz* غراتس, *Greenwich* غرينتش, *Gregorius* غريغوريوس, *Grenada* غرينادا, *Groningen* غرونينغن, *Guadalajara* غوادالاخارا, *Guam* غوام, *Guangzhou* غوانزو, *Guayaquil* غواياكيل, *Guyana* غيانا, *Hamburg* هامبورغ, *Hamburgini* هامبرغيني, *Hong Kong* هونغ كونغ, *Hugo* هوغو, *Hungary* هنغاريا, *Johannesburg* جوهانسبرغ, *Lady Gaga* غاغا , *Lagos* لاغوس, *Leningrad* لينينغراد , *Logan* لوغان, *Luxembourg* لوكسمبورغ, *Magnolia* ماغنوليا, *Margaret* مارغريت, *Margarita* مرغريتا, *Megan* ميغان, *Ortagus* اورتاغوس, *Prague* براغ, *Red Tag* رد تاغ, *Viet Cong* فيت كونغ, *wang* وانغ, *Young* يونغ, *Gabon* الغابون, *Guatemala* غواتيمالا, *Togo* توغو, *Shanghai* شنغهاي; *Gürkan Şef* غوركان شيف, *Nescafe Dolce Gusto* نسكافيه دولتشي غوستو, *Bangladesh* بنغلاديش, *Guilin* غويلين, *Haemoglobin* هيموغلوبين, *Mega, Mall* ميغا مول, *Megaغ*ميغا, *Megawatt* ميغاوات, *Morgan*مورغان, *Spring* سبرينغ . In these examples, orthography and pronunciation diverge as many Arabic speakers override the spelling and pronounce غ as /g/, especially in educated or international contexts. These examples reflect a broader transliteration compromise: غ is used as a placeholder, but phonetic intuition corrects it in speech. This distinction is crucial for educators, translators, and AI systems. It shows that transliteration is not just about matching letters, but it is about respecting phonetic origin, dialectal norms, and the semantic domain.

Moreover, in 37% of nouns in the whole sample, the Arabic equivalents are transcribed with ج. Here, there are 2 subsets. In the first subset (27%), ج *is pronounced /g/* as in the following examples: *American Eagle Outfitters* أميريكان إيجل أوتفترز, *Blogger*بلوجر, *Blooming* بلومينج , *Burger King*برجر كينج, *Copenhagen* كوبنهاجن, *Craig* كريج, *Gallery* جاليري , *Galvus met* جالفس مت جالفوس, *Gap* المحل جاب, *Garden* جاردن, *Gateau* جاتوه, *Gigabyte* جيجابايت, *Gloria Jean* جلوريا جين, *Godiva* جوديفا, *Golf* جولف, *Goody* جودي, *Google*جوجل, *Gordon* جوردون, *Grand Store* جراند ستور, *Greenland* جرينلاند, *Grill* جريل, *Group*جروب , *Guess* المحل جس, *Gurkan Chef* جوركان شيف, *Mango* مانجو , *moringa* مورينجا, *Mr Games*مستر جيمز, *Mugabi* موجابي, *Nuggets* ناجتس , *Oregano* أوريجانو, *Penang* بينانج, *Reading (city or act)* ريدينج, *Spring roll* السبرينج رول, *St Petersburg* سان بطرسبرج, *Suprise Gift* سربرايز جفت, *Tully Gore* تولي جور. In the second subset (10%), ج is pronounced /d͡ʒ/ especially by Arabic speakers who have poor knowledge of English or those from countries like Sudan, Jordan and Palestine as in the following examples: *Anglo-Saxon*انجلو ساكسوني, *Colgate* كولجيت, *Cigarettes* السجائر, *Galaxy* جالكسي, *Gallon* جالون, *Glucophage* جلوكوفاج, *Glucose* جلوكوز, *Glycogen*جلايكوجين , *graphics* جرافيكس, *guanti* جوانتي, *Gyroscope* الجيروسكوب, *England* إنجلترا, *English* إنجليزي, *Glycerine*جلسرين, *Maggie* ماجي.

Additionally, only one noun was transliterated with ق by MC (*hashtag* هاشتاق) and another one by GT (*Guilin*قويلين), and another one was transliterated with ک (*Malacca* ملكا instead of ملقا).

It is noteworthy to say that both MC and GT gave a single Arabic transliteration to each English noun. On the contrary, in human transliterations, especially in Saudi Arabic, there are several Arabic examples transliterated with ق as *Goody* قودي, *Gurkan Chef* قوركان شيف, *Oregano* اورقانو, *Golf* قولف, *Nescafe Dolce Gusto* نسكافيه دولس قوستو, *Grand Store* قراند ستور, *Mr Games* مستر قيمز, *Guzal* قوزال للملابس التركية, *grammar* قرامر, *Reading* ريدينق. Another 16% of the nouns in the sample have two or more human transliterations such as:

- *Google* كوكل غوغل گوگل چوچل جوجل قوقل جوجل
- *Group* جروب، قروب، گرو غروب
- *England* إنكلترا أو إنكلترة انجلترا
- *English* الإنجليزية الأنقليزية الانكليزية; *English* انجليزي انقليزي
- *Hashtag* هاشتاج هاشتاق هاشتاغ
- *Blogger* بلوجر بلوغر بلوقر
- *Golf:* غولف جولف
- *In Burger King* برجر كنج & برغر كينج، بيرغر, *gas* جاز غاز , *gigabyte* جيجابايت غيغابايت, *glycerine* الغليسيرين جليسرين, *Greenland* غرينلاندا جرينلاندا, *grill* جريل غريل, *haemoglobin* هيموجلوبين هيموغلوبين, *Maggie* ماجي ماغي, *mega mall* ميجا مول, ميغا مول , *mega* ميغا ميجا, *megawatt* ميغاواط ميجاوات, *Morgan* مورجان مورغان, *the two Arabic variants have* ج & غ.

4. Discussion

This study investigated how English and foreign words with g are transliterated in Arabic, and why transliteration variations exist. It was found that Arabic speakers transliterate /g/ with equivalents containing غ، ج، ق، ک, whereas AI transliterates them with either ج or غ only. These transliteration variations arise from a complex interplay of several factors. The first factor is regional and dialectal. In human transliteration, the closest Arabic sound to /g/ is influenced by the speaker's Arabic dialect. For instance, in Egyptian Arabic, /g/ is transliterated by ج, as all words spelled with ج in spoken Egyptian Arabic are pronounced with /g/ (جمال /gama:/, جميل /gami:l/, جنة /ganna/). But in Levantine Arabic, /g/ is transliterated by غ (*Gracia* غراسيا) and names of many cities and countries. Interestingly, in Egyptian Arabic, *Ghana* غانا, *Guinea* غينيا , *gas* غاز; Chicago شيكاغو are transliterated with غ not ج for Arabization and historical reasons. In Saudi Arabic, the replacement of /g/ sound by ق is common as all words that are written with ق are pronounced with a /g/ as in قال /ga:l/ (said), قلم /galam/ (pen), قمر /gamar/ (moon), قلب /galb/ (heart) and so on. Saudi college students started to use ق in transliterating English words with /g/ (*English* انقليزي, *hashtag* هاشتاق, *group* قروب, *grammar* قرامر, *reading* ريدنق) a couple of decades ago with the prevalence of colloquial language and colloquial spelling on social media. This was not the case three decades ago as Arabic speakers were using Standard Arabic than the vernacular in written communication.

Secondly, the transliteration of /g/ using ق reflects lack of knowledge on the part of the students and non-specialists especially in the age of social media. Students' lack of knowledge of the English and Arabic phonological systems, leads to basic errors. Many Arabic speakers, on social media nowadays, ignore Standard Arabic spelling, spell the words the way they pronounce them in their local dialect (Al-Jarf, 2023). This is also supported by studies by Haider et al. (2025) and Awang and Salman (2017) which touch upon how regional preferences influence transliteration choices for borrowed and technical terms.

Thirdly, some names in the geopolitical domain have old, established transliterations (Chicago شيكاغو) that are widely recognized in all Arab countries. If such names were transliterated with ج, Arabic speakers who do not know English, would decode *Singapore, Afghanistan, Ghana, Guinea* …etc with ج /dʒ/, which sounds cacophonous compared to pronouncing then with /ɣ/, i.e., غ is preferred to avoid /dʒ/ confusion. Al-Omar (2013) discussed this tension between "fact" (established form) and "implication" (the accurate form). Many of the city and country names spelled with غ, as mentioned earlier (Singapore سنغافورة, Afghanistan أفغانستان , Ghana غانا , Guinea غينيا, Chicago شيكاغو or gas غاز) are Arabized and pronounced in spoken Arabic with غ, so they are spelled with غ to match their pronunciation.

A fourth explanation is lack of standardization and policy. Unlike languages with official language academies that publish binding rules such as the French Academy, there is no single authority whose transliteration rules are followed across the Arab World. Arabic Language Academies in Cairo, Amman and Damascus lack authority that mandates the use of certain forms. Different institutions (e.g., academic journals, media outlets, governments) use different conventions. Lack of standardization and policy which creates confusion, was confirmed by Al-Omar (2013) and Muscat (2011).

Moreover, some contextual and pragmatic factors and the influence of purpose and audience affect transliteration variations in Arabic. For example, a transliteration for a scholarly article may aim for phonetic precision, while a transliteration for a brand (e.g., "Microsoft") may prioritize consistency and recognition over perfect phonetics. A transliteration on a street sign for tourists will differ from one in a local government document (Gu, 2025).

Regarding the transliteration of English nouns containing /g/, MC and GT gave accurate Arabic transliterations of 98% of the nouns in the sample, with two major transliteration variants: one set with غ, (61%) and the other one with ج (37). The percentage of correct transliterations rendered by MC and GT are higher than the correct translation of medical terms (68.6% by MC and 74.5% by GT) (Al-Jarf, 2024); Arabic expressions of Impossibility (52% by MC) (Al-Jarf, 2025c); Gaza-Israel war terminology (48% by MC and GT) (Al-Jarf, 2025a); Arabic grammatical terms used metaphorically (29% by DeepSeek, 23.5% by GT and 43% by Copilot) (Al-Jarf, 2025b); zero-expressions (29% by MC and GT) (Al-Jarf, 2025d); and translation of technical terms by GT (Al-Jarf, 2021a; Al-Jarf, 2016).

To decide between غ and ج, ق in transliterating English names with /g/, AI systems often rely on pronunciation dictionaries like the CMU Pronouncing Dictionary (CMUdict), which maps English words to their phonetic representations using ARPAbet symbols. AI also takes into consideration dialectal variations. For example, in Gulf Arabic: ج /ʤ/, غ or ق may be used for /g/. In Egyptian Arabic, ج /g/ is preferred. In Maghrebi Arabic, transliterators may use گ (non-standard letter for /g/). AI may adapt transliteration based on target dialect or regional usage patterns. In addition to CMUdict, AI tools/assistants use the following resources: (i) AMR-KELEG Transphonator for Rule-based phoneme-to-Arabic transliteration; (ii) Yamli for Smart Arabic keyboard that learns transliteration patterns; (iii) ContentTool Transliterator - an online transliteration tool with adaptive logic; (iv) Social Media & Wikipedia for Real-world transliteration usage (e.g., قوقل, جرافيكس, غابون). AI systems are also trained on large Arabic-English corpora from Wikipedia, news archives, social media and so on and learn transliteration patterns from real-world usage. If 90% of the sources write "Google" as قوقل, AI will follow suit.

Furthermore, AI takes into account the semantic domain. Formal, geopolitical terms often use غ (e.g., غابون for Gabon); technology, consumer and pop culture often use ج (e.g., جالاكسي for Galaxy); and informal Saudi usage may use ق (e.g., قوقل for Google). Personal and consumer names with hard g are transliterated with ج. These are often western, commercial, or tech-related. Arabic uses ج to reflect branding familiarity, even if it risks phonetic ambiguity. Furthermore, ج is used in tech/media (*Graphics* جرافيكس), in tech branding (*Galaxy* جالكسي), for legacy transliteration (*Anglo-Saxon* أنجلو ساكسوني), in food branding (*Burger* برجر), in consumer terms (*Cigarettes* سجائر), for brand name (*Colgate* كولجيت, for European place name (*Copenhagen* كوبنهاجن), for colonial legacy term (*England* إنجلترا), in media/arts (*Gallery* جاليري), in measurement terms (*Gallon* جالون) and in logos, packaging, and tech. Egyptian Arabic pronounces ج as /g/, influencing regional transliteration as mentioned above.

5. Recommendations and Conclusion

To transliterate names from English to Arabic accurately, some prior studies emphasized linguistic and phonotactic rules and the importance of understanding the sound systems (phonology) and writing conventions of both English and Arabic to create accurate and natural-sounding transliterations. In this respect, Haider et al. (2025) recommended a translation-oriented approach that considers whether a phonetic (sound-based) or orthographic (spelling-based) strategy is more appropriate for a specific word and context, acknowledging that both have a place. Alshuwaier & Areshey (2011) specifically recommended using phonotactic rules (rules about how sounds can be combined in a language). Their approach involves converting English text to a phonetic representation (like IPA) and then mapping those sounds to their closest Arabic equivalents while adhering to Arabic syllable structure constraints. Odisho (1992) provided a foundational linguistic recommendation stating that transliteration should be primarily phonemic (sound-based) rather than graphemic (letter-based). It emphasizes the need for a systematic one-to-one correspondence between the sound systems of English and Arabic to minimize distortion. Aziz (1983) proposed detailed grapheme-to-grapheme (letter-to-letter) rules for transliterating proper nouns. This is a more traditional rule-based approach that establishes direct correlations between English letters and Arabic characters.

Other prior studies proposed solutions for the transliteration inconsistencies through standardization, education, and official policy. These solutions include: (i) standardization and Arabicization policies from Arabic Language Academies to ensure technical terms are transliterated consistently across the Arab world (Awang & Salman, 2017). (ii) Arabic Language Academies and media outlets should agree upon and adhere to a unified set of rules for transliterating Proper Nouns to avoid confusion (Al-Omar, 2013). (iii) National governments should establish official, standardized transliterations for place names (toponyms) to be used on maps, in official documents, and for tourism (Muscat, 2011). (iv) EFL teachers should explicitly teach standardized transliteration rules to students to overcome common errors and create consistency in academic work (Ali & Ra'uf, 2010).

Further prior studies highlighted contextual and pragmatic factors. These studies argued that accuracy is not just technical; it is also about how the transliteration is received and used in a specific context. They commended that transliteration choices must consider sociolinguistic factors, identity, and intended audience (e.g., tourists vs. locals), not just linguistic accuracy (Gu, 2025). Purpose and audience must be considered. A transliteration for a scholarly text might prioritize precision, while one for a popular magazine might prioritize common usage and recognition (Haider et al. 2025; Al-Omar (2013). An official state standardization is recommended as a pragmatic solution to a real-world problem of navigation and official identity (Muscat, 2011).

Moreover, prior studies like Abd El-Wahab et al. (2022) and [25] Asar and Abdullah (2025) recommended the use of advanced technology to overcome historical challenges, especially in new digital contexts. Since human-created rules are insufficient for the vast variability in language, some prior studies proposed Machine Learning (ML) and Statistical Models, hybrid and data-driven approaches that combine the best rule-based and statistical methods, using real-world data to improve transliteration accuracy such as smart, standardized digital tools and APIs that can provide consistent transliteration for apps, OS, and social media platforms to overcome current fragmentation (Asar & Abdullah, 2025); advanced machine learning models (like Neural Networks or Sequence-to-Sequence models) that can automatically learn complex grapheme and phoneme mapping relationships between English and Arabic, moving beyond older statistical methods for higher accuracy (Abd El-Wahab, Abu-Khzam & El Den, 2022); a hybrid framework that combines probability-based models with web mining & scrapes the web to find common transliteration variants used in real-world contexts (e.g., online news) to inform and validate the statistical model (Zhou, Huang & Chen, 2008); methods for automatically creating massive sets of transliteration pairs from parallel corpora and web data, which are essential for training accurate statistical and ML models Fattah & Ren, 2008); a statistical, generative approach based on building a model of "how English names become Arabic names." This involves learning joint probability distributions of letter/phoneme sequences from a large bilingual corpus of names, making it highly effective for Information Retrieval tasks (AbdulJaleel & Larkey, 2003); leveraging parallel corpora (large collections of aligned text in both English and Arabic) and using a system that can learn reliable, contextually-appropriate mappings by analyzing how names are consistently transliterated in professionally translated texts (e.g., news archives) (Samy, Moreno & Guirao, 2005).

Arabic speakers should use غ or ج only in transliterating English words containing g, as these are common in all Arabi countries, whereas گ, ڭ, ڤ , چ and ق are used in some regions and are not wide spread in Arab countries. Many Arabic speakers in most Arab countries are not familiar with them and will decode them phonetically which will lead to inaccurate pronunciation. Similarly, speakers of Saudi Arabic should not use ق for words pronounced with g in English as قروب قوقل هاشتاق as other speakers of other Arabic dialects will read قوقل قروب as /qru:p/ and /qu:qil/ which sound awkward and meaningless.

Furthermore, translation students and social media users can use AI tools and Assistance such as MC, GT, Gemini DeepSeek and others to transliterate personal, city, country, shop, product, brand, hotel names and other nouns borrowed from English to Arabic as they transliterate them with high accuracy. They should specify in the instructions that they need a Standard Arabic transliteration.

Since issues about transliterating English sounds/letters that do not exist in Arabic and Arabic sounds/letters that do not exist in English have not been extensively studied, especially in the age of Artificial Intelligence, this study recommends that future researchers explore how place names and proper nouns mentioned in the Holy Quran are transliterated by AI.

Conflicts of Interest: The author declares no conflict of interest.

ORCID ID: https://orcid.org/0000-0002-6255-1305

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Abd El-Wahab, M., Abu-Khzam, F. & El Den, J. (2022). *An effective machine learning approach for English-Arabic transliteration*. 4th International Conference on Natural Language Processing (ICNLP), 345-349. IEEE.

[2] AbdulJaleel, N., & Larkey, L. (2003). English to Arabic transliteration for information retrieval: A statistical approach. *Center for Intelligent Information Retrieval Computer Science, University of Massachusetts*.

[3] Ali, Z. & Ra'uf, M. (2010). Some problems of Iraqi EFL students in transliteration. *Journal of Al-Qadisiya University, 13*(3), 7.

[4] Al-Jarf, R. (2025a). AI translation of the Gaza-Israel war terminology . *International Journal of Linguistics, Literature and Translation*, *8*(2), 139-152. https://doi.org/10.32996/ijllt.2025.8.2.17. Google Scholar

[5] Al-Jarf , R. (2025b). DeepSeek, Google Translate and Copilot's translation of Arabic grammatical terms used metaphorically. *Journal of Computer Science and Technology Studies*, *7*(3), 46-57. DOI: 10.32996/jcsts.2025.7.3.6. Google Scholar

[6] Al-Jarf, R. (2025c). 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. https://doi.org/10.32996/jcsts.2025.7.8.33. Google Scholar

[7] Al-Jarf, R. (2025d). Translation of Zero-Expressions by Microsoft Copilot and Google Translate. *Journal of Computer Science and Technology Studies (JCSTS)*, *7*(2), 203-216. https://doi.org/10.32996/jcsts.2025.7.2.20. Google Scholar

[8] Al-Jarf, R. (2024). Translation of medical terms by AI: A comparative linguistic study of Microsoft Copilot and Google Translate. I2COMSAPP'2024 Conference, In Y. M. Elhadj et al. (Eds.): I2COMSAPP 2024, LNNS 862, pp. 1–16, 2024. Springer Nature, Switzerland AG. https://doi.org/10.1007/978-3-031-71429-0_17. Google Scholar

[9] Al-Jarf, R. (2023a). Absence of vowels in the English spelling of Arabic personal names on social media. *International Journal of English Language Studies (IJELS), 5*(4), 88-97. Doi: 10.32996/ijels.2023.5.4.7. ERIC ED633828. Googe Scholar

[10] Al-Jarf, R. (2023b). English spelling of Arabic compound personal names by educated Arabs on Facebook. Journal of Humanities and Social Sciences Studies (JHSCS), 5, 1, 53-64. DOI: 10.32996/jhsss.2023.5.1.8. Google Scholar

[11] Al-Jarf, R. (2023c). English spelling of the glottal stop and voiced pharyngeal fricative in Arabic personal names by educated Arabs on Facebook. International Journal of English Language Studies, 5(1), 11-22. https://doi.org/10.32996/ijels.2023.5.1.2. Google Scholar

[12] Al-Jarf, R. (2023d). Non-conventional spelling in informal, colloquial Arabic writing on Facebook. International Journal of Linguistics, Literature and Translation, 6, 3, 35-47. DOI: 10.32996/ijllt.2023.6.4.6. Google Scholar

[13] Al-Jarf, R. (2023e). Semantic and syntactic anomalies of Arabic-transliterated compound shop names in Saudi Arabia. *International Journal of Arts and Humanities Studies* (IJAHS), 3(1), 1-8. DOI: 10.32996/Ijahs.2023.3.1.1. Google Scholar

[14] Al-Jarf, R. (2022a). Deviant Arabic transliterations of foreign shop names in Saudi Arabia and decoding problems among shoppers. *International Journal of Asian and African Studies (IJAAS), 1*(1), 17-30. DOI: 10.32996/ijaas.2022.1.1.3. Google Scholar

[15] Al-Jarf, R. (2022b). English transliteration of Arabic personal names with the definite article {al-} on Facebook. *British Journal of Applied Linguistics (BJAL)*, *2*(2), 23-37. DOI: 10.31926/but.pcs.2022.64.15.2.2. Google Scholar

[16] Al-Jarf, R. (2022c). Gemination errors in Arabic-English transliteration of personal names on Facebook. *International Journal of Linguistics Studies (IJLS)*, *2*(2), 163-170. Google Scholar

[17] Al-Jarf, R. (2022d). Variant transliterations of the same Arabic personal names on Facebook. *International Journal of English Language Studies (IJELS)*, *4*(4), 79-90. DOI: 10.32996/ijls.2022.2.2.18. Google Scholar

[18] Al-Jarf, R. (2021a). An Investigation of Google's English-Arabic Translation of Technical Terms. *Eurasian Arabic Studies, 14*, 16-37. Google Scholar

[19] Al-Jarf, R. (2021b). Linguistic-cultural characteristics of hotel names in Saudi Arabia: The case of Makkah, Madinah and Riyadh hotels. *International Journal of Linguistics, Literature and Translation (IJLLT), 4*(8), 160-170. DOI: 10.32996/ijllt.2021.4.8.23. Google Scholar

[20] Al-Jarf, R. (2016). *Issues in translating English technical terms to Arabic by Google Translate*. 3rd International Conference on Information and Communication Technologies for Education and Training (TICET 2016), pp. 17-31. Khartoum, Sudan, March 12-14. Google Scholar

[21] Al-Jarf, R. (2015). *English and Arabic writing systems for translation students.* Academia.edu . Google Scholar

[22] Al-Jarf, R. (2003). *Contrastive phonology*. King Saud University. Retrieved from https://1filedownload.com/wp-content/uploads/2020/01/Contrastive-Phonology-Transparencies.pdf. Google Scholar

[23] Al-Jarf, R. (1995). *An Arabic word identification diagnostic test for the first three grades*. Center for Educational Research. College of Education. King Saud University. Google Scholar

[24] Al-Jarf, R. (1990). *English and Arabic phonology for translation students*. King Saud University. Google Scholar https://www.researchgate.net/publication/281003181. Google Scholar

[25] Al-Jurf, R. S. (2002). A contrastive analysis of English and Arabic for translation students. King Saud University. Google Scholar

[26] Al-Jurf, R. (1994). *Contrastive phonetics for translation students*. King Saud University. Google Scholar https://www.researchgate.net/publication/281003427. Google Scholar

[27] Al-Omar, N. (2013). Transliterating proper nouns: facts and implications. *Linguistics, Culture and Education*, *2*(2), 119-132.

[28] Alshuwaier, F. & Areshey, A. (2011). *Translating English names to Arabic using phonotactic rules*. In Proceedings of the 25th Pacific Asia Conference on Language, Information and Computation, 485-492.

[29] Asar, A. & Abdullah, A. (2025). Arabic transliteration for digital smart communication: Challenges and solution. *4th International Conference on Computing and Information Technology (ICCIT)* (pp. 647-651). IEEE.

[30] Awang, R. & Salman, G. (2017). Translation and Arabicization methods of English scientific and technical terms into Arabic. *AWEJ for Translation & Literary Studies*, *1*, 2.

[31] Aziz, Y. (1983). Transliteration of English proper nouns into Arabic. *Meta*, *28*(1), 70-84.

[32] Fattah, M. & Ren, F. (2008). English-Arabic proper-noun transliteration-pairs creation. *Journal of the American Society for Information Science and Technology, 5*9(10), 1675-1687.

[33] Gu, C. (2025). *Multiscriptal English in transliterated linguistic landscapes*. Cambridge University Press.

[34] Haider, A., Saed, H., Albarakati, M., Abu Tair, S. & Jarrah, S. (2025). Phonetic and orthographic transliteration of borrowed words in Arabic: A translation-oriented study. *WORD*, *71*(1), 1-19.

[35] Muscat, O. (2011). The English transliteration of place names in Oman. *Journal of Academic and Applied Studies*, *1*(3), 1-27.

[36] Odisho, E. (1992). Transliterating English in Arabic. *Zeitschrift für arabische Linguistik*, (24), 21-34.

[37] Samy, D., Moreno, A. & Guirao, J. (2005). *A proposal for an Arabic named entity tagger leveraging a parallel corpus.* In International Conference RANLP, Borovets, Bulgaria, 459-465.

[38] Zhou, Y., Huang, F., & Chen, H. (2008). Combining probability models and web mining models: a framework for proper name transliteration. *Information Technology and Management*, *9*(2), 91-103.