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

Copilot and Gemini's Translation of English and Arabic "Half"-Based Expressions

Reima Al-Jarf
Full Professor of English and Translation Studies, Riyadh, Saudi Arabia
Corresponding Author: Reima Al-Jarf, **E-mail:** reima.al.jarf@gmail.com

| ABSTRACT
This study compares the translation accuracy of two AI models: Microsoft Copilot (MC) and Google Gemini (GG), in translating English and Arabic half-expressions, the direction that is easier, types of translation errors they commit and the translation strategies they utilize. A sample of 100 Arabic and 100 English half-expressions was randomly selected from *Al-Maany Dictionaries* and Google searches, representing both general-language and specialized-domains. Results showed that in the Arabic half-expressions set, MC and GG produced 80% correct equivalents, consisting of 61% identical equivalents in form and meaning and 19% correct equivalents with variant wording. Additionally, GG supplied 12% extra correct equivalents not produced by MC, raising its total accuracy to 92%, compared to 80% for MC. In the English half-expressions set, MC and GG produced 56% correct equivalents, including 21% identical equivalents in form and meaning and 35% correct equivalents with variant wording. GG provided 24% additional correct equivalents, while MC supplied 9% additional correct equivalents. Overall, GG achieved 80% accuracy, compared to 65% for MC. Error analysis revealed consistent MC and GG failures in culture-specific expressions, social and colloquial Arabic expressions, commercial-domain terminology (furniture, fabrics, vehicles), expressions with polysemous words, word-order errors, and cases where GG produced mixed correct and incorrect equivalents. The most common translation strategies utilized by MC and GG included literal translation, explanation (specifying the domain, context in which the expression is used, and a definition), giving multiple equivalents, and English equivalents containing *semi*, *hemi*, *mid*, *middle*, and *demi*. Transliteration was the least frequent. Arabic-to-English translation was easier than English-to-Arabic as reflected by a higher percentage of correct equivalents and differences in lexical density and synonymy. Compared with the author's prior studies, half-expressions were significantly easier for AI to translate than zero-expressions, Abu/Umm-brand names, animal and plant names, and Gaza–Israel war terms, and Arabic grammatical terms used metaphorically. Explanations are provided for the highly identical equivalents given by MC and GG, why Arabic-English translation was easier than English-Arabi translation, and why GG outperformed MC. Recommendations for AI developers and translation students are given as well.

| KEYWORDS
Half-expressions, fixed and formulaic expressions, English-Arabic language pairs, AI translation, Copilot, Gemini, literal translation, contextual translation, machine translation evaluation, blind testing.

| ARTICLE INFORMATION

ACCEPTED: 15 June 2026 **PUBLISHED:** 14 August 2026 **DOI:** 10.32996/jcsts.2026.8.8.25

1. Introduction

According to the Cambridge Dictionary, "half" denotes one of two equal parts of a whole, a period in a sports game, or a state of partiality or incompleteness. Grammatically, it functions as a noun (cut an apple in half), an adjective (half an hour, half brother), and an adverb expressing partial states (half asleep). Furthermore, the conceptual domain of "half" in English extends beyond the standalone lexeme to include productive morphological prefixes and combining forms such as *semi-*, *hemi-*, *demi-*, *mid-*, *middle*, *bi-*, and *di-*, which similarly involve partiality, halving, or midpoint positioning across general and specialized expressions. In everyday English, "half" and its related forms frequently convey the meaning of "partly" or "somewhat," and appear in idiomatic expressions that highlight compromise, insufficiency, exaggeration, or hesitation. For instance, compromise is encoded in expressions such as *half a loaf is better than none*, *meet someone halfway*, and *split the difference*, while incompletion and inadequacy surface in *half-cocked*, *half-baked*, *do things by halves*, and *half-hearted*. Exaggeration and emphasis appear in *not half bad*, *half dead*, *given half a chance*, and *too clever by half*, whereas identity and relational expressions emerge in *my other half*,

better half, and *half the battle*. Cognitive and psychological states are likewise expressed through *half a mind*, *believe half of what you see*, and *in half a heartbeat*. Together, these uses show how "half" and its morphological counterparts, function as a conceptual marker for partiality, hesitation, insufficiency, and relational complementarity.

In Standard Arabic, the equivalent term نصف and the Colloquial نص exhibit similar semantic richness as English. The plural أنصاف (halves), the verbs نصف (split into 2 halves) and انتصف (to reach half of something) and مناصفة (sharing half of a thing) reflect the same conceptual domain of dividing, halving, or reaching a midpoint. These derived forms parallel the English patterns of partiality and proportional division, demonstrating that "half-based" expressions are prevalent across both specialized and everyday registers.

Arabic-نِصْف (half) also belongs to a broader lexical field that includes synonyms such as صَدْر, صَدْع, صِدْع, شَطْر, شَقّ, شِقّ, صَدع, فِلْقَة, مِصْراع (Al-Maany). Although these words differ in nuance, they all share the semantic feature of partiality, division, or incompleteness. This semantic property allows them to function as attenuating devices in discourse, similar to English expressions like "partly," "halfway," or "in part." When used metaphorically or evaluatively, these Arabic terms soften the force of the statement, reduce commitment, and create mitigated or attenuated expressions. Their presence in fixed language highlights the role of lexical partiality in shaping pragmatic tone and in signaling reduced intensity within Arabic lexical and idiomatic structures.

In both English and Arabic, "half", نِصْف and نُصّ expressions function as attenuating devices that soften the semantic force of a statement. They signal partiality, reduced intensity, or weakened commitment. Although they are not metaphors by themselves, they play a central role in forming attenuated expressions within Arabic discourse and frequently appear in mitigated metaphors, evaluative descriptions, and pragmatic softeners.

Beyond common usage, *half* & نصف expressions are highly productive across general and specialized disciplines. Scientists rely on *half* to conceptualize decay, reduction, thresholds, partial effectiveness, or incomplete states, and these lexical extensions often evolve into technical or semi-technical terminology. In physics, the notion appears in "half-life"; in chemistry, in "half-reaction"; in biology and medicine, in "half-maximal effective concentration (EC50"); in engineering and computer science, in "half-duplex communication"; in environmental science, in the "half-saturation constant"; and in economics and social sciences, in the "half-life of policy effects". Even mathematics uses "half" metaphorically in expressions such as نصف قطر (radius). Across these fields, half serves as a conceptual tool for describing processes that are transitional, partial, or evolving, revealing how language often draws on everyday cognitive schemas to articulate complex phenomena.

Due to their prevalence in general and specialized use, different kinds of expressions, metaphors, proverbs and idioms have been the focus of a substantial body of research investigating how AI systems translate different types of metaphors. Several studies focused on AI handles culturally loaded fixed expressions, idioms and proverbs. These include a corpus-based study of idiom translation by AI (Mughal et al., 2024); Additionally, literary metaphors have received considerable attention, as seen in metaphors in literary machine translation (MT) (Karakanta, Nas & Dorst, 2025), assessing MT of emotional depth, metaphorical complexity, and cultural nuances in literary texts (Fang & Wang, 2025), a comparative analysis of literary translation by AI with a focus on Poe's The Tell-Tale Heart (Iskakbayeva et al., 2025), and a perspective on literary metaphor in the context of generative AI (Van Heerden & Bas, 2024).

Cultural metaphors were explored in studies such as translating culture-specific expressions by AI systems (Azizov, 2024), a conceptual synthesis of cognitive semiotics and AI pragmatics (Vance-Kaelen, 2026), and AI recognition and misinterpretation of culturally embedded symbolic language (Assiri, 2026).

A fourth group of studies examined metaphor translation in specialized domains. Political metaphors were analyzed in cultural perspectives as in the translation of political text metaphors using AI research (He & Jiang, 2024); the role of AI in the translation of metaphors in tourism discourse (Spinzi, 2025); a critical evaluation of AI translation of conceptual metaphors of light and darkness in the Qur'an (Farshi & Zare, 2025); a survey on computational metaphor processing techniques (Ge, et al., 2023); corpus and metrics for evaluating MT quality of metaphorical language (Wang et al., 2024); assessing novelty in AI-generated metaphors (Bassenge, 2026); AI translation of metaphors for the EFL classroom (Brala Vukanović, 2025); general metaphor interpreting challenges by AI (Skrynnikova, 2024) and AI and the burden of bad metaphors (Kūlis, 2026). These studies emphasize the inherent difficulty AI systems face when processing figurative language.

Another group of studies focused on specific language pairs and how AI performance varies across linguistic and cultural contexts. English-related research includes on metaphor translation into English based on AI (Wang & Chai, 2024), translation of metaphors in Ex Machina from English to Indonesian (Nafiza & Erwina, 2025), exploring AI's capability in translating English metaphors into Arabic (Alazzam et al., 2025), assessment of English-Russian MT systems in translating conceptual metaphors (Erokhin, 2026), AI

translation of English metaphors in financial and economic reports translated into Spanish and French (Serón Ordóñez, 2026). Chinese-related studies include exploring the translation of idioms by AI in Chinese internet literature (Guo & Amirdabbaghian, 2025) and a cognitive stylistic approach to the translation of spatial metaphors in Alai's Chen'ai Luoding (Yu & Shao, 2023). Additional language-specific investigations include Indonesian (Nafiza & Erwina, 2025), Russian (Erokhin, 2026), Macedonian (Shalevska et al., 2025), Spanish and French (Serón Ordóñez, 2026), and Yoruba (Raji-Ellams, 2025) translating Macedonian nonliteral proverbs and idioms by AI (Shalevska et al., 2025), cross-linguistic interpretation of metaphor in Yoruba proverbs (Raji-Ellams, 2025).

Arabic-focused studies form an important subset, including translating Arabic literary metaphors by ChatGPT-4.0 and Google Translate 2024 (Ahmed, Dewi & Anis, 2025); AI-powered translation of Arabic idioms into English (Al Salem et al., 2026); exploring AI's capability in translating English metaphors into Arabic (Alazzam et al., 2025); translation of conceptual metaphors of light and darkness in the Qur'an by AI models (Farshi & Zare, 2025); AI-powered translation of Arabic idioms into English (Al Salem, Abusalim & Alghazo, 2026); AI-translated idioms and proverbs between Arabic and English (Darwish, et al. (2025). Together, these studies demonstrate the multilingual and cross-cultural nature of metaphor translation and highlight the complexity AI systems face when navigating figurative language across diverse linguistic pairs.

Additionally, Al-Jarf's research (2012–2025) provides a comprehensive examination of how AI language models handle Arabic metaphorical, idiomatic, and culturally-embedded expressions. Her studies include Copilot's English translation of contrastive emphatic negation (Al-Jarf, 2025f), and ESP students' use of AI for translating chemical nomenclature (Al-Jarf, 2025e). She also investigated AI translation of English and Arabic "sleep" terms and formulaic expressions (Al-Jarf, 2025m; Al-Jarf, 2005n); denotative and metonymic abu- and umm- animal and plant folk names (Al-Jarf, 2025g), abu-brand names (Al-Jarf, 2025d), and folk medical terms with Om and Abu (Al-Jarf, 2025k), and translation of medical terms by Copilot and Google Translate (Al-Jarf, 2024a & Al-Jarf, 2024a). Additional studies examined AI translation of Arabic expressions of impossibility (Al-Jarf, 2025l), grammatical terms used metaphorically (Al-Jarf, 2025h), zero-expressions (Al-Jarf, 2025o), educational polysemes in full-text research articles (Al-Jarf, 2025a), Gaza–Israel war terminology (Al-Jarf, 2025b), and medical terminology (Al-Jarf, 2024a, 2024b), and human vs AI translation of chemical compound names (Al-Jarf, 2025j). Earlier work explored Google's translation of technical terms (Al-Jarf, 2021; 2016) and multilingual translation performance in 2012 and 2025 (Al-Jarf, 2025i). She also analyzed calligraphic text recognition by Gemini, Ernie Vilg, and Google Translate (Al-Jarf, 2025c). Collectively, these studies reveal persistent challenges in AI translation of Arabic metaphors, idioms, folk naming systems, and culturally-loaded expressions, highlighting the linguistic depth required for accurate interpretation.

The above literature review reveals a scarcity of studies examining how AI language models process and translate fixed and formulaic expressions containing the English element "half" and Arabic "نصف". While previous research has extensively explored AI performance on idioms, zero-expressions, and folk naming systems, the specific linguistic behavior, error patterns, and comparative accuracy of models like Microsoft Copilot (MC) and Google Gemini (GG) regarding half-based expressions remain unexplored. This study aims to address the gap in AI translation research by examining how two AI models—MC and GG—render English and Arabic half-expressions. Specifically, it seeks to: (i) evaluate and compare the overall translation accuracy and the types of equivalence (identical vs. variant wording) produced by MC and GG; (ii) identify the major sources of error and recurrent failure patterns in the translation of half-based expressions; (iii) analyze the translation strategies employed by both models; (iv) compare AI performance on half-expressions with the author's prior findings on other metaphorical expression types; (v) determine which translation direction is easier for AI—Arabic-English or English-Arabic; (vi) explain why MC and GG produced more correct equivalents in Arabic-English than in English-Arabic translation; (vii) examine why MC and GG generated identical equivalents in several cases; (viii) identify which model outperformed the other; (ix) outline the implications of the findings for AI-based translation of half-expressions; and (x) highlight the limitations of current AI models in handling half-expressions.

This study is significant because it provides the first systematic examination of how AI language models interpret fixed and semi-literal expressions containing the element "half" (and "نصف") across English and Arabic, a domain that has been largely overlooked in current research. By analyzing the performance of leading AI models: MC and GG, the study highlights the linguistic challenges these systems face when processing such expressions. The findings contribute to a deeper understanding of bilingual lexical patterns within computational contexts, offer a new framework for evaluating AI translation accuracy, and provide practical insights for improving AI-based translation, interpretation, and educational tools. This research also supports the development of more accurate AI systems for Arabic and English users.

2. Methodology

A corpus of 200 English half-expressions and 300 Arabic half-expressions (containing "نصف" or "نُصّ") was collected from Al-Maany English–Arabic and Arabic–Arabic dictionaries and Google Search. Both sets included expressions used in general contexts as well as a wide range of specialized domains, including math, computer science, dentistry, textiles, seamanship,

telecommunications, engineering, physics, nuclear physics, electricity, electronics, mechanics, plumbing, botany, music, military terminology, printing, skiing, boxing, sports, tennis, and rugby. From this corpus, two randomized samples of 100 English and 100 Arabic half-expressions were selected to ensure manageable coding, data handling, and statistical analysis.

The half-expressions were presented to MC and GG in two batches: 100 Arabic expressions followed by 100 English expressions. They were provided in isolation and evaluated through blind, exhaustive-equivalent prompts that did not specify the type of expression nor the domain in which it is used. By presenting the expressions without contextual scaffolding (e.g., verbs, subjects, or surrounding text), all external cues were removed, forcing MC and GG to rely solely on internal token-level associations and static lexical representations. This design made it possible to observe how well each model handles multi-word noun phrases and idiomatic expressions when deprived of contextual support.

Each Arabic half-expression was translated into English, and each English half-expression was translated into Arabic by both MC and GG. Each model was prompted to provide all possible equivalents it could generate. They were instructed to write each equivalent next to the source expression and not to alter the order of items in either list.

Two master tables were created—one for the Arabic set and one for the English set—containing columns for the source expressions, MC output, evaluation and coding fields, and GG output. For each expression, the models' responses were entered side-by-side. Coding began by evaluating and comparing the equivalents produced by MC and GG to determine whether they represented identical correct responses, correct responses with different wording, cases where only one model produced a correct equivalent, or cases where one or both models produced faulty equivalents. Items sharing the same analytical outcome were color-coded for consistency and ease of comparison. All coding, data analysis, and statistical procedures were conducted manually and solely by the author.

Percentages were calculated for: correct identical equivalents produced by both MC and GG, correct equivalents with different wording produced by both models, correct equivalents produced by MC only, and correct equivalents produced by GG only.

Translation errors were classified into the following categories: failure to recognize culture-specific expressions; pragmatic mismatches in social labeling; misinterpretation of social/layperson expressions; commercial domain errors (furniture, fabrics, vehicles); polysemy-related errors; word-order errors (MC); word-order and structural inversions; institutional and administrative contextualization; and mixed responses (GG).

The translation strategies used by MC and GG were categorized into: literal translation; explicitation; contextual integration; domain explanations and pragmatic enrichment; domain-specific contextual integration; explanation (domain, context, and definition); double and multiple equivalents; use of semi-, hemi-, mid-, middle-, and demi-based equivalents; communicative and idiomatic equivalence; and transliteration.

Percentages of errors in each category were calculated separately for MC and GG. Both quantitative and qualitative analyses were conducted to interpret the models' performance patterns.

For reliability and validity, two colleagues specializing in translation and linguistics independently classified a sample of errors representing semantic, contextual, syntactic, and lexical weaknesses. They reviewed the list of half-expressions and their equivalents and evaluated the accuracy and classification of each translation. Classifications by all three evaluators were compared, yielding 97% inter-rater agreement, with disagreements resolved through discussion.

3. Results

3.1 Arabic Corpus Performance (100 items)

Data analysis of the Arabic half-expressions showed that MC and GG gave correct responses to 80% of the items. Correct responses include 61% identical correct equivalents (in meaning and form) and 19% correct equivalents with the same meaning but a different wording (See Tables 1 & 2 for examples). In addition, GG gave additional correct equivalents to 12% of the items. This means that GG gave correct equivalents to a total of 92% compared to a total of 80% correct equivalents by MC. GG gave faulty responses to 8% of the items while MC gave 12% faulty responses (See Table 3 for examples).

Table 1: Examples of Identical Correct Equivalents in Arabic-English and English-Arabic Translation By MC and GG

English-Arabic	Arabic -English
• half-hearted فاتر / بلا حماس • half marathon نصف ماراثون • half mast منتصف السارية (تنكيس العلم)	• إِنْتَصَفَ الطَّرِيقَ He reached the halfway point of the road • التَّدْبِيرُ نِصْفُ الْمَعِيشَةِ Management is half of livelihood (Frugality is half of wealth)

• half Nelson مسكة نصف نيلسون (في المصارعة) • half ration نصف حصة غذائية • half reaction نصف تفاعل (في الكيمياء الكهربائية) • half roll نصف لفة • other half النصف الآخر (الزوج/الشريك) • second half الشوط الثاني • top half النصف العلوي • word half نصف كلمة	• It loses half of its value تفقد نصف قيمتها • He split the prize with his colleague تَنَصَّفَ الجائِزَةَ مَعَ زَميلِهِ • Hemiplegia / Paraplegia (Partial paralysis)شلل نصفيّ • Focal radius length طول نصف القطر البؤري • Half-integer عدد نصف صحيح • Midterm break عطلة نصف الفصل الدراسي • Half-grill restaurant = half a chicken مطعم نصف شواية • Half a loss is better than a total loss نِصْفُ الْخَسَارَةِ وَلاَ الْخَسَارَةُ كُلُّهَا • Half a point نصف نقطة

Table 2: Examples of Correct Equivalents Given by MC and GG Semantic/Syntactic Variations in Arabic-English Translation

Arabic half-expressions	Copilot (MC)	Gemini (GG)
• نصف الشهر	**Mid**-month	**Middle** of the month
• نصف الملعب	Half the field	Half of the field / Half-court
• نِصْفُ التُّفَّاحَةِ	Half an apple	Half of the apple
• نصف الحقيقة	Half the truth	Half-truth
• نصف العمى ولا اعمى كله	Half-blind is better than fully blind	Better half-blind than totally blind
• النّصف الحلو	The sweet half (women)	The better half (spouse)
• نصف المستقيم أو الشعاع	A half-line / ray	Ray / **Semi**-infinite line
• قانون محيط نصف الدائرة	**Semi**circle circumference formula	**Semi**circle perimeter formula
• قماش نصف وقفه كوري	Half-weight Korean fabric	Korean **semi**-stiff fabric
• نصف طقم سوبر	Half a super set	Super half-set (jewelry)
• نصف حل	A half-solution	Half-measure / Partial solution
• نصف دورة	Half-turn / half-rotation	**Semi**-circle / Half-turn / Half-cycle
• نصف ناضج	Half-cooked / medium-done	Half-ripe / Medium-well (food)
• نصف الموسم	Half the season	**Mid**-season
• نصف قطر الزاوية	Angle radius	Angular radius
• نصف قطر مدار الكوكب	Planetary orbital radius	Orbital radius of the planet

Table 3: Examples of Faulty Equivalents Given by MC and GT in Arabic-English Translation

	MC	GG
رجالٌ نِصْفٌ	Half-men	• Middle-aged men
رجلٌ نِصْفٌ	A half-man	• A middle-aged man
رنة خلخال نصف تولة	The chime of an anklet weighing half a tola	• Half a Tola (traditional weight unit)
نص انصيص	Half of a half	• Exactly half / Precisely divided
مفرد و نصف	One and a half	• Single and a half (size/capacity)
حوض نصف ركبة	Half-knee basin	• Semi-pedestal sink / Half-kneeling basin
حركات نصف كم	Half-hearted movements	• Cheap tricks/Shady or unprofessional behavior
نص مستوي	A flat half	--
نص نص	Half of a half	--
نصف كم	Half a kilometre	--

3.2 English Corpus Performance (100 items)

As for the English half-expressions set, data analysis revealed that MC and GG gave correct responses to 56% of the items. This includes 21% identical correct equivalents in meaning and wording and 35% correct equivalents with the same meaning but different wording. This means that GG gave more correct equivalents (80%) compared to 65% correct equivalents by MC. In addition, GG, but not MC, gave correct equivalents to an additional 24% of the items that were not produced by MC, while MC supplied an additional 9% correct equivalents not rendered by GG. Both models gave faulty responses to 5% of the items, GG gave faulty responses to an additional 9% faulty responses and MC gave an additional 24% faulty responses (See Tables 4, 5, 6, & 7).

The models performed better on the Arabic-to-English set than the English-to-Arabic set. GG's performance on both sets was higher compared to MC. GG achieved 92% vs 80% on the Arabic-English translation task, whereas GG achieved 80% vs 65% by MC on the English-Arabic translation task. Both GG & MC gave identical translations for 60% of the Arabic and 48% of the English expressions. (شلل نصفيّ, half speed, half Nelson). Arabic-to-English expressions were easier to translate than English-to-Arabic

expressions, reflecting differences in lexical density and synonymy. Both GG & MC failed with Arabic heritage (folklore) نص انصيص, social slang نصف كم, and commercial metaphors (رنة خلخال نصف تولة). AI models very frequently tend to translate literally (word for word). GG & MC rendered equivalents that have semantic & syntactic (wording) variations (See Tables 4).

Table 4: Examples of Variant Equivalents in English – Arabic Translation by MC and GG

	Copilot	Gemini
half wit	أبله / محدود الذكاء	معتوه / ضعيف العقل
half witted	غبي / ناقص الذكاء	ناقص العقل / أبله
half size	نصف الحجم / نصف المقاس	نصف الحجم
subtracter half	نصف المُطرح (في الدارات المنطقية)	نصف طارح
right half	النصف الأيمن	لاعب الجناح الأيمن (قديم)
half-proof	نصف قوة (للمشروبات الكحولية)	نصف العيار (للكحول)
half-tongue	نصف لسان/لسان ناقص (لغويًا أو تشريحيًا)	لسان لثغ (أو لغة هجينة)
Not half	ليس قليلًا / ليس نصفًا (تعبير للمبالغة)	جداً / ليس قليلاً على الإطلاق
half time	استراحة منتصف المباراة / نصف الوقت	الاستراحة (بين الشوطين)

Table 5: Examples of Faulty Equivalents Given by Both MC and GT in English-Arabic Translations

	Copilot	Gemini
Half adder	جامع نصف	نصف جامع (في الدوائر الإلكترونية)
half hardy	نصف قوي التحمل	شبه متحمل للبرودة (في النباتات)
half lap	وصلة نصف تعشيق (في النجارة)	نصف لفة / وصلة تراكبية
half leather	نصف جلد	تجليد نصف جلدي
Half right	نصف يمين / انحراف نصف يمين	نصف مقطع (رسم هندسي)
half router	قاطع نصف تعشيق (في النجارة)	نصف محق / نصف يمين
Half section	مقطع نصفي (في الرسم الهندسي)	نصف مقطع (رسم هندسي)
half size	نصف الحجم / نصف المقاس	نصف الحجم
half life	نصف العمر	عمر النصف (في الفيزياء والفيزياء النووية)

Table 6: Examples of Faulty Equivalents Given by MC Only in English-Arabic Translation

• halfcock (في الأسلحة) نصف سحب	• halfbound (noun): binder, bookbinder تجليد نصف / مجلّد
• Halfcoupling وصلة نصف	• half seas over ثمل نصفه
• halfevergreen نصف دائم الخضرة	• halftitle صفحة نصف العنوان
• half rest استراحة نصف	• halfvalue layer = halfvalue thickness طبقة نصف القيمة
• Half right نصف يمين	• half cock نصف سحب
• half smash ضربة نصف سماش	• half cut نصف مقطوع
• half title نصف عنوان	• halfproof نصف إثبات / نصف تركيز
• half volley (تنس) نصف فولي	• half and half clasp / half and half clasps مشبك نصف بالنصف / مشابك نصف بالنصف
• half subtracter (دائرة منطقية) مطروح نصف	• Half term نصف الفصل الدراسي
• duplex half نصف مزدوج	• half timer عامل نصف وقت
• halfbinding تجليد نصف	• half router مِسحَج نصف
• halfblooded نصف الدم / نصف الأصل	• subtracter half نصف مطروح
• halfmeasure إجراء نصف / حل غير كامل	

Table 7: Examples of Faulty Equivalents Given by GT Only in English-Arabic Translation

• scrum half: (في الرغبي) لاعب نصف السكروم	• half shadow: (الظليل) شبه الظل
• half adjust: التقريب لأقرب نصف / ضبط تقريبي بالنصف	• half broad: قماش نصف عريض / نصف واسع
• half cast: مولود من نژادين / هجين العرق	• half tide: منتصف وقت المد والجزر
• Half Duplex: شبه مزدوج (نظام إرسال واستقبال غير متزامن)	• half back: لاعب نصف الخلف

3.3 Types of Errors Made by MC and GG

Failure to Recognize Culture-Specific Expressions

Both MC and GG failed to interpret several culture-specific expressions that require shared folk memory. For instance, neither model was able to understand the popular expression نص انصيص MC rendered it as half of a half and GG as exactly half, instead of preserving the name as Nas Nusayis. In Levantine folklore, Nas Nusayis is a fictional character symbolizing cleverness,

resourcefulness, and the ability to overcome difficulties despite being thin and small in size. A similar issue appears in رنة خلخال نصف تولة for which MC produced the chime of an anklet weighing half a tola and GG gave half a tola, without recognizing that Rannat Khulkhal is the name of a women's perfume from Taif. These examples show that language models require heritage-based and commercial context to correctly interpret traditional cultural markers.

Pragmatic Mismatches in Social Labeling

A recurring source of error, particularly for MC, is pragmatic mismatch when translating social labels and descriptors. The models often failed to account for the social nuance inherent in these terms, leading to "over-literalization" that results in offensive or awkward English equivalents. For example**:**

- أنصاف الرجال**:** MC rendered this as "Half-men (weak men)," failing to grasp the idiomatic weight. GG, however, identified the social context: "Half-men (cowards or those lacking manhood)".
- أنصاف مجانين: While MC provided a literal "Half-mad people," GG accurately identified the idiomatic descriptor: "Half-wits".

Misinterpretation of Social/Layperson Expressions

The models also misinterpreted expressions referring to social categories. MC translated نصف رجال / نصف امرأة / نصف رجل as a half woman, while GG rendered it as a middle-aged woman. The correct meaning, however, is layperson or average man/woman—a socially neutral term referring to an ordinary, average individual. Such expressions cannot be interpreted literally because they belong to popular social discourse rather than formal linguistic structure.

Commercial Domain Errors (Furniture, Fabrics, Vehicles)

Commercial terminology posed a major challenge for both models. MC translated مفرد ونصف as one and a half instead of queen bed, because the expression is not mathematical but a Gulf market label where "مفرد ونصف" corresponds to a queen-size bed. MC also rendered قماش نصف وقفه كوري as half weight Korean fabric rather than semi-stiff Korean fabric, reflecting limited familiarity with popular textile terminology. In the automotive domain, both MC and GG translated نصف نقل and سيارة نص نقل as half load transport, semi-truck, or half truck, whereas the correct meaning is pickup truck. In plumbing, MC translated حوض نصف ركبة as half knee basin instead of semi-pedestal basin, revealing a lack of idiomatic knowledge in this technical field.

Polysemy-Related Errors

Polysemy was a major source of error. GG translated نصف دورة عطوفية as semi-affectionate revolution and half affectionate cycle due to the multiple meanings of the word عطوفية, whereas the correct interpretation is semi-rotational cycle. MC likewise rendered نص نص as half of a half instead of so-so or fifty-fifty. In إِنْتَصَفَ النَّهارُ, MC produced midday arrived rather than it is midday / it is noon. For حركات نصف كم, MC gave the literal "half-hearted movements" instead of the idiomatic meaning "cheap tricks or unprofessional behavior". In نص مستوي, MC translated the phrase as "flat half" instead of "medium cooked or semi-flat". MC also misinterpreted نصف كم (without diacritics) as "half a kilometer" rather than "short sleeve", because the lack of diacritics and contextual cues prevented the model from selecting the correct meaning.

Word-Order Errors (MC)

MC also produced faulty syntactic ordering in several expressions. For example, it translated نصف طقم سوبر as half a super set instead of super half set. This type of error reflects a weakness in nominal element ordering when handling Arabic commercial constructions.

In 4% of the English expressions, GG produced equivalents with faulty word order, compared to 3% in MC.

- Examples of faulty word order by GG**:** half leather**:** تجليد نصف جلدي, half hardy**:** شبه متحمل للبرودة (plants), half life**:** عمر النصف (physics), half track**:** شاحنة مجنزرة نصفياً.
- Examples of faulty word order by MC**:** half hardy نصف قوي التحمل, half life نصف العمر, half adderنصف جامع.

Word-Order and Structural Inversions

Both models occasionally struggled with the syntactic restructuring of complex English-Arabic compound terms, leading to minor structural inversions or awkward phrasing (*Word-Order Errors*). This shows their intermittent difficulty in maintaining standard specialized nomenclature order. For example**:**

- half life**:** Both models introduced a structural shift, producing (العمر النصفي) instead of the standard scientific convention (عمر النصف في الفيزياء والفيزياء النووية).
- Half adder**:** MC produced a faulty word-order inversion (جامع نصف), failing to align with the standard electronics terminology (نصف جامع في الدوائر الإلكترونية).
- half hardy**:** Both models exhibited structural or syntactic weaknesses, rendering it with awkward phrasing like (نصف قوي التحمل) instead of the proper botanical equivalent (شه متحمل للبرودة في النباتات).

Institutional and Administrative Contextualization

The models displayed noticeable performance gaps when translating institutional, labor, or socio-economic terms that require an understanding of specific operational contexts rather than a compositional translation. For example:

- *half timer:* MC failed to capture the institutional context and produced a faulty equivalent, whereas GG accurately translated it based on labor/academic systems (عامل أو طالب بنصف وقت).
- *half router*: MC generated a confusing literal translation (مِسحَج نصف), while GG provided the proper technical equivalent (موجه نصف شبكي / راوتر جزئي).
- *half board / half broad*: MC struggled with commercial and textile terminology variations, whereas GG successfully separated structural layouts from general linguistic composition.

Mixed Responses (GG)

GG frequently produced mixed outputs combining correct and incorrect equivalents. For instance, in نصف نقلit offered semi truck, half truck, and light pickup truck, with only the last one being correct. In نصف غشائي /غشائي النصف, it generated semi membranous, half membranous, semi membrane, and membrane half, several of which are not terminologically acceptable. A similar blend of technical and literal interpretations appeared in GG's rendering of half track as شاحنة مجنزرة نصفياً. MC likewise produced inconsistent outputs, translating half time as استراحة منتصف المباراةland نصف الوقت.

3.4 Strategies Used by MC and GG

Literal translation

MC gave more literal translations in the faulty responses than

The translation data analysis showed that both MC and GG frequently resorted to literal translation as a default strategy whenever cultural, commercial, or polysemous "half-expressions" appeared. As demonstrated in the results section, literal translation emerged whenever contextual or domain-specific knowledge was required, leading to inaccurate or misleading equivalents. Furthermore, both MC and GG relied on literal translation even in cases where their output happened to be correct—whether fully equivalent or differently phrased. This is because many of the expressions in the dataset are technical, commercial, or structurally similar across Arabic and English, which makes literal rendering appear accurate on the surface. In other words, the correctness of some outputs does not indicate successful semantic interpretation; rather, it reflects a coincidence where the literal form aligns with the English expression. Thus, literal translation emerges not only in erroneous cases but also as a general default strategy whenever the models encounter specialized terminology or expressions with cross-linguistic structural similarity.

Both models exhibited a tendency toward literal translation when handling precise quantitative or technical measures. This strategy often proved faulty when the Arabic expression functioned as a specialized nomenclature rather than a simple mathematical fraction. For example: for نصف العشر, MC defaulted to a literal "Half of a tithe," which is mathematically vague, whereas GG successfully mapped it to the specific financial/legal value "One-twentieth (5%)". Similarly, for نص مستوي, MC produced "A flat half," which is semantically null, whereas GG employed domain knowledge to provide "Medium-cooked / Semi-flat".

Explicitation, Contextual Integration, Domain Explanations and Pragmatic Enrichment

This strategy was employed primarily by GG to resolve cultural or domain-specific ambiguities. Rather than providing a blunt literal translation, GG often appended context to ensure semantic clarity for the target audience. For example:

- نصف الوقت *(Half-time):* GG clarified the domain as "(sports)," whereas MC provided only the base term.
- مطعم نصف شواية*:* GG translated this as "Half-grill restaurant (referring to half a chicken)," bridging the gap between the literal Arabic label and the intended culinary meaning.
- نص ذبيحة*:* GG provided the necessary domain context as in "Half a carcass (butchery term)".

Domain-Specific Contextual Integration

Compared to MC, GG showed superior capability to integrate domain-specific context directly into translations. While MC frequently defaults to ambiguous or literal renderings of specialized terminology, GG incorporates precise contextual markers across various fields (such as mechanics, military, and publishing) to ensure target-readability. For example:

- *half cock*: MC produced a faulty literal translation (نصف سحب), whereas GG correctly integrated the military/weaponry context: *(*وضع النصف حميّة / تأهب غير مكتمل للمطرقة في السلاح*)*.
- *Half coupling*: MC failed with a direct literal output (وصلة نصف), while GG provided the exact technical domain mapping: *(*وصلة نصفية في السباكة والميكانيكا*)*.
- *half title*: MC rendered it literally as (نصف عنوان), whereas GG successfully incorporated the publishing context: *(* عنوان فرعي / عنوان داخلي مختصر في الطباعة*)*.

Explanation (domain, context and definition)

GG frequently used domain-based explanation, providing the field in which the expression is used, as opposed to MC. In 31% of the items, GG supplied the domain in which the half-expressions is used compared to 6% by MC, as in Table below.

• half-closed interval: (**الرياضيات** في) فترة نصف مغلقة, • half open interval; half open interval: (**الرياضيات** في) فترة نصف مفتوحة, • half angle and half side formulae of spherical trigonometry: **المثلثات الكروية** في الضلع ونصف الزاويه ونصف صيغ, • half veneer crown / half veneer crowns: تاج غشائي نصفائي (**في طب الأسنان**), • half cap crown / half cap crowns: تاج نصف غطائي / تيجان (**في طب الأسنان**) نصف غطائية, • Half Kaddish: (**صلاة في اليهودية**) كاديش نصف كامل, • half cup: (**في الملابس**) نصف كوب / مقاس كوب نصفائي (**الداخلية**), • half raglan: (**في تفصيل الملابس**) كم نصف راجلان, • half astern: (**في الملاحة البحرية**) نصف سرعة إلى الخلف, • half deck: (**في السفن**) سطح جزئي / نصف سطح, • Half word: (**وحدة بيانات حاسوبية**) نصف كلمة, • half call: (**في الاتصالات**) مكالمة مجهزة جزئياً / نداء نصف مكتمل • Half section: (**في الرسم الهندسي**) نصف مقطع, • half-life: (**في الفيزياء والفيزياء النووية**) عمر النصف,	• half coil: (**في الكهرباء**) نصف ملف / نصف لفة, • Half adder: (**في الدوائر الإلكترونية**) نصف جامع, • Half coupling: (**في السباكة والميكانيكا**) وصلة نصفية, • half evergreen: (**للنباتات**) **شبه** دائم الخضرة, • half hardy: (**في النباتات**) **شبه** متحمل للبرودة. • half note: (**في الموسيقى**) نوتة بيضاء / نصف نوتة زمنية, • half rest: (**في الموسيقى**) سكتة نصفية (نصف زمن سكتة في), • Half right: (**في العسكرية**) نصف محق / محق إلى حد ما (وفي لليمين در جزئياً, • half cock: وضع النصف حميّة (تأهب غير مكتمل للمطرقة في **السلاح**), • half title: (**في الطباعة**) عنوان فرعي / عنوان داخلي مختصر, • half pipe: (**منحدر للتزلج**) أنبوب نصف دائري, • right half: (**في الرياضات الجماعية**) لاعب وسط أيمن, • half nelson: (**في المصارعة**) مسكة نصف نيلسون, • half smash: (**في كرة الريشة/المضرب**) ضربة ساحقة متدرجة/جزئية, • half volley: (**في التنس/الكرة**) ضربة مرتدة مباشرة بعد الملامسة, • scrum half: (**في الرغبي**) **السكروم** لاعب نصف,

In 39% of the English expressions, GG gave the context and gave an explanation as opposed to 3% by MC of the equivalent. The following are examples:

• half arsed: (**عامية**) مهمل / مقتضب / منجز بغير إتقان • half pint: (**عامية**) نصف باينت / شخص صغير القامة • half cut: (**عامية**) ثمل / سكران جزئياً • by half: بالنصف / أكثر مما ينبغي (تُستخدم **للتأكيد** مثل: *too clever by half*) • half boot: (**تصل لمنصف الساق**) جزمة قصيرة • half breed; half breed: **مولود من عرقين** / هجين النسب **مختلفين** • half flood; half tide: منتصف مد البحر / منتصف وقت المد والجزر • half-holiday; half-holiday: **عطلة بعد** نصف يوم عطلة (**الظهيرة**) • half-staff; half-staff (v): (**العلَم**) ينكس • half thickness layer (= half value layer): طبقة نصف السمك (**سمك القيمة النصفية**) • half eagle: النصف نسر **الذهبية** **عملة** • half thickness: طبقة نصف السمك (سمك النصف الممتص) • half tongue: محاكمة بلغة مشتركة/ترجمان / **يتحدث بلغة غير متقنة** • Not half: ليس بالقليل أبداً / تماماً وشدة (**تعبير مجازي** يعني "كثيراً جداً") • half baked: غير ناضج / غير مدروس / فكرة غير مكتملة • half dozen; half dozen: نصف درزن / ستة قطع • half eagle: نصف نسر (**عملة ذهبية أمريكية قديمة بقيمة 5 دولارات**) • Half fare: نصف أجرة / تذكرة بنصف السعر • half mooned: مقوس كالهلال / **على شكل نصف قمر**	• half-truth: (**حقيقة منقوصة ومضللة**) نصف حقيقة • half baked: غير ناضج / فكرة فطيرة غير مدروسة • halfhearted: فاتر / فاقد للفتور والحماس • half maternal aunt/uncle: (**من طرف الأم فقط**) خالة/خال غير شقيق • half binding: (**تجليد الكتب بجلد على الجوانب والظهر فقط**) تجليد نصف جلدي • half-blooded: هجين / من سلالتين / أخ أو اخت غير شقيقة • half bound: binder, bookbinder: كتاب مجلد تجليداً نصفياً (مُجَلِّد) • half measure: إجراء ناقص / تدبير نصف مكتمل • half seas over; half sea over: (**تعبير مجازي قديم**) ثمل / مخمور • half title: (**مقدمة الكتاب** في) صفحة العنوان المختصر • half value layer = half value thickness: سمك القيمة النصفية (**سمك المادة اللازم لتقليل الإشعاع إلى النصف**) • half cock: وضع عدم الجاهزية التامة / وضع النصف حمية • half proof: درجة كحول نصف قياسية / إثبات جزئي • half and half clasp / half and half clasps: مشبك/**مشابك الأسنان النصفية المتوازنة** • half adjust: التقريب لأقرب نصف / ضبط تقريبي بالنصف • half cast: مولود من نژادين / هجين العرق • half-track: (**مركبة نصف مجنزرة**) شاحنة مجنزرة نصفياً • duplex half: (**شبه** مزدوج) اتجاه واحد في الاتصال المزدوج • Half Duplex: (**نظام إرسال واستقبال غير متزامن شبه** مزدوج) • half shadow: (الظليل) **شبه** الظل

Double & Multiple Equivalents

GG produced double or multiple equivalents more frequently than MC. In the Arabic set, GG gave multiple equivalents for 34% of the items compared to 29% by MC. In the English set, GG provided multiple equivalents for 44% of the items, whereas MC did so for 13% (See Table 6 for examples).

Table 6: Examples of Double and Multiple Equivalents Given by MC and GG

Copilot	Gemini
• نصف متبلور Semi-crystalline / Half-crystalline • نِصْفُ الشَّفَّافِيَّة Half-transparency / Semi-transparency / Partial transparency • نِصف إهليلَجي Semi-elliptical / Half-elliptical • نصف واعي Semi-conscious / Half-conscious • نصف محمل Half-loaded / Semi-loaded • نصف مشحون Half-charged / Semi-charged • نصف نَفوذ Semi-permeable / Half-permeable • نصف منفِذ Semi-penetrating / Partially penetrative • نصف ناضج Half-cooked / medium-done • نصف المستقيم أو الشعاع A half-line / ray • نِصف مَحدود Semi-limited / Partially constrained • نصف حلقي، نصف إطاري Semi-annular / Semi-circular frame / Half-ring / Semi-frame • نِصف رُبيعي Semi-quarterly / Half-quarter (rare) • مهلك النصف Half-lethal / Semi-lethal / Half-destructive • نظرية نصف Half-theory / Semi-theory • تقليدية نصف كلاسيكية Semi-classical traditional / Half-classical traditional • وتري النصف Semi-chordal / Half-chord • نصف الأخدود Half-groove / Semi-groove • شلل نصفيّ Hemiplegia (paralysis of one side) • مناصفة Equally / fifty-fifty • منتصف الطريق Midway / halfway • نصف الكلام Half-speech / incomplete speech • نصف سري Semi-secret / Half-secret / Partially confidential • نصف عام Semi-annual / Half-yearly • نصف وتري النسيج Semi-chordal texture / Half-chord textile pattern • نصف غشائي النصف؛ غشائي semi-membranous, half-membranous, semi-membrane • نصف نقل Half-load transport / Semi-transport capacity • سيارة نص نقل Half-transport vehicle / Semi-transport truck • امرأَةٌ نِصْفٌ A half-woman (incomplete woman) • واحدة النصف Half-unit / One-half unit • نصف دورة عطوفية semi-rotational cycle, half-rotational cycle	• نص مستوي Medium-cooked / Semi-flat • حركات نصف كم Cheap tricks / Shady or unprofessional behavior • نصف متبلور Semi-crystalline, Half-crystalline • نِصْفُ الشَّفَّافِيَّة Semi-transparency, Demi-transparency, Half-transparency • نِصف إهليلَجي Semi-elliptical, Semi-elliptic, Half-elliptical • نصف واعي Semi-conscious, Half-conscious, Semi-aware • نصف محمل، نصف مشحون Semi-loaded, Half-loaded, Semi-charged, Half-charged • نصف المستقيم أو الشعاعRay / Semi-infinite line • نِصف رُبيعي Semi-quadrature, Half-quarterly, Semi-quartile • مهلك النصف Half-destroyer, Destroyer of half, Semi-destructive • نظرية نصف، تقليدية نصف كلاسيكية Half-theory, Semi-classical, Semi-traditional, Semi-classical conventional • وتري النصف Semi-chordal, Semi-fibrous, Half-chordal, • نصف الأخدود Semi-grooved, Half-groove, Semi-sulcate • شلل نصفيّ Hemiplegia / Paraplegia (Partial paralysis) • مناصفة Equally / Fifty-fifty / Halves • منتصف الطريق Halfway / Midpoint • نصف الكلام Half of the conversation / Incomplete speech • نصف سري Semi-secret, Quasi-secret, Half-secret • نصف عام Semi-public, Semi-general, Half-public • نصف غشائي، غشائي النصف Semi-membranous, Semi-membrane, Half-membranous, Membrane-half • نصف نقل Semi-truck, Half-truck, Light pickup truck • سيارة نص نقل Semi-pickup truck, Half-truck vehicle, Pickup truck • واحدة النصف Unit-and-a-half, One-and-a-half unit Half-unit • نص عمر Half-life / Used / Middle-aged • حوض نصف ركبة Semi-pedestal sink / Half-kneeling basin • نص نص So-so / Fifty-fifty • إنْتَصَفَ النَّهارُ It became noon / midday

Giving Equivalents with Semi, Hemi, Mid, Middle, Demi

A further strategy observed in both MC and GG is the use of English prefixes and combining forms such as *semi-*, *hemi-*, *mid-*, and *middle* when rendering equivalents to Arabic expressions containing نصف. In the Arabic dataset, MC produced equivalents with *semi* in 40% of the items, compared to 55% by GG, most of which appeared in the identical and correct equivalents (see Table 6). In 8% of the items, both models used *hemi*, particularly in expressions referring to anatomical or spatial halves, such as *hemiplegia* and *hemisphere*. GG additionally employed *mid* and *middle* in 12% of the Arabic items, compared to 7% by MC, producing equivalents such as *middle-aged man*, *midday*, *mid-season*, *mid-month*, *middle of the month*, *Mid Sha'bān*, *midpoint*, *midway*, and *midterm*.

In addition to these productive prefixes, both MC and GG occasionally utilized the prefix *demi-*. Although this pattern was rare (2% in the Arabic set, and 1% in the English set), the models produced select equivalents such as *demigod* and *demi-transparency*. The use of *demi* reflects the models' sensitivity to English lexical items that encode partiality through inherited or less common morphological forms. Unlike *semi* and *hemi*, which are productive and widely used in technical terminology, *demi* appears primarily in fixed lexical items. Its occurrence therefore indicates that the models rely not only on productive morphological rules but also on stored lexical knowledge when mapping Arabic expressions denoting "half" to English.

This entire pattern reflects the models' sophisticated reliance on cross-linguistic morphological alignment. When an Arabic expression denotes a spatial, temporal, or quantitative "half," the models systematically map it to appropriate English equivalents. The variation between *semi-* , *hemi-* , and *mid-/middle* stems from differences in English morphological conventions and lexical domains rather than semantic flaws, showing that the models can successfully navigate multiple standardized morphological devices in English.

Communicative and Idiomatic Equivalence

GG consistently prioritized communicative equivalence, aiming to capture the functional meaning of Arabic proverbs and metaphors rather than their literal components. For example**:**

- التدبير نصف المعيشة**:** GG captured the idiomatic essence**:** "Management is half of livelihood (Frugality is half of wealth)," whereas MC merely provided a direct translation.
- النصف الحلو**:** GG successfully mapped this to the English cultural equivalent**:** "The better half (spouse)," while MC opted for a less idiomatic rendering**:** "The sweet half (women)".

Transliteration

Transliteration was a rare strategy employed by both MC and GG when encountering specialized sports or technical terms that are often borrowed or partially nativized in Arabic discourse. Instead of providing a fully semantic or descriptive equivalent, the models occasionally resorted to phonetic transcription. For instance, GG transliterated *scrum-half* as "مركز السكرّم هاف" (in rugby) instead of a descriptive functional equivalent like "لاعب وسط التشكيلة". Similarly, MC transliterated *half smash* as "ضربة نصف سماش" instead of "ضربة ساحقة متدرجة/جزئية" (in badminton/tennis), and rendered *half volley* as "نصف فولي" (in tennis) rather than " ضربة مرتدة مباشرة بعد الملامسة". This rare strategy indicates that when faced with specialized jargon containing "half," the models occasionally bypass translation in favor of direct phonetic transfer, reflecting the influence of loan-word usage in sports journalism.

4. Discussion

4.1 Why MC and GG produced more correct equivalents in Arabic–English than English-Arabic Translation

The Arabic-to-English direction was easier for both models because the Arabic half-expressions in the dataset were standard, frequent, and possessed direct English equivalents, establishing a direct one-to-one correspondence in form and meaning. Google Gemini (GG) produced only 8% faulty responses, while Microsoft Copilot (MC) produced 20%. Difficult items were confined to a small subset of expressions drawn from colloquial Arabic, culturally embedded folk terminology, and popular market language. These include expressions such as حوض نصف ركبة, رنة خلخال نصف تولة, قماش نصف وقفة كوري, and نص انصيص. These terms carry socio-cultural, commercial, or dialectal meanings that cannot be recovered through literal decoding. While widely used in everyday contexts involving clothing, fabrics, bedding, jewelry, and marketplace terminology, they lack fixed or standardized English equivalents. In contrast, the Arabic half-expressions translated into English were semantically transparent with stable, exact English counterparts, making them straightforward for both human learners and AI models. This asymmetry reflects not inherent linguistic complexity, but rather the fact that the Arabic items evaluated were familiar and directly mappable, whereas the corresponding English items often required idiomatic, cultural, and folk-based interpretation that MC is less equipped to perform.

This directional asymmetry is further amplified by the structural bias of large language models toward English. As previous research has shown, these models are trained on substantially larger and more diverse English corpora than Arabic corpora, resulting in stronger English lexical, idiomatic, and domain coverage. Copilot in particular tends to default to literal translation as its baseline strategy—a behavior consistent with earlier findings in the author's studies comparing Arabic and English metaphors and idioms. Those prior studies identified four cross-linguistic patterns: (i) Expressions identical in form and meaning in both languages (ii) Expressions identical in meaning but different in form. (iii) Expressions existing only in English. (iv) Expressions existing only in Arabic. The first category—identical form and meaning—is universally the easiest for both human learners and AI models, and most Arabic half-expressions in the present dataset fall neatly into this category. As attenuated metaphors, they lack overt figurative structure and carry transparent meanings, unlike other metaphor types (such as grammatical metaphors like أصبح في خبر كان) that require deep cultural and syntactic inference.

By contrast, the difficult items encountered in the dataset belong to colloquial usage, culturally embedded folk terms, and commercial expressions common in regional markets. These expressions do not have one-to-one equivalents and require cultural

decoding rather than literal mapping. Consequently, Arabic-to-English translation proved easier because the tested Arabic items were familiar, frequent, and possessed exact English counterparts, whereas the reverse direction demanded complex idiomatic, cultural, and domain-specific interpretation.

Finally, certain Arabic half-expressions derive entirely from folk culture rather than from standard lexical or morphological patterns. For example, نص انصيص is not a compositional phrase but the name of a fictional character from traditional children's stories passed down in popular culture. Because such expressions originate from oral folklore and are deeply culturally embedded rather than semantically transparent, AI models cannot infer their meanings through morphology or syntax alone. This explains why both students and language models struggle with these specific items: they demand deep cultural grounding and narrative context rather than algorithmic text processing.

4.2 Why MC and GG Gave Identical Equivalents (Model Convergence and Divergence)

Results showed that MC and GG gave correct responses to 80% of the items in the Arabic set. This includes 61% identical correct equivalents and 19% correct equivalents with a different wording. In the English half-expressions set, results revealed that MC and GG gave correct responses to 56% of the items. This includes 21% identical correct equivalents and 35% correct equivalents with a different wording.

The Arabic-to-English direction produced a much higher percentage of identical correct equivalents (61%) compared to the English half-expressions (21%). This asymmetry reflects fundamental differences in how each language encodes "half" and how predictable the mapping is from Arabic to English. This asymmetry reflects fundamental differences in how each language encodes "half" and how predictable the mapping is. Arabic expressions containing نصف are generally semantically explicit, structurally transparent, and typically denote a single, unambiguous meaning. Because English has standardized morphological devices for expressing partiality—such as *semi-*, *hemi-*, *mid-*, and *middle*—the cross-linguistic mapping from Arabic to English is highly stable. As a result, both MC and GG tend to converge on the same standardized English equivalent, producing identical correct outputs at a much higher rate.

In contrast, English half-expressions do not map onto Arabic with the same level of uniformity. Arabic lacks a single standardized system for expressing "half," and the language naturally allows multiple valid renderings for the same concept. A single English item may correspond to نصف, شبه, منتصف, وسط, or a derived form, depending on register, region, or stylistic preference. This structural and lexical richness means that even when MC and GG produce semantically correct Arabic equivalents, they often choose *different* correct forms rather than identical ones. Consequently, the percentage of identical equivalents in the English → Arabic direction is much lower.

Furthermore, many English half-expressions are polysemous or context-dependent. An English item may encode spatial, temporal, quantitative, or metaphorical meanings, and Arabic offers several ways to express each of these meanings. This inherent flexibility increases variation in the Arabic outputs, even when they are semantically accurate. By contrast, the Arabic items in the dataset are largely technical and non-metaphorical, which minimizes ambiguity and leads both models to select the same English equivalent.

Taken together, these linguistic and terminological factors explain why identical correct equivalents are substantially higher in the Arabic set (61%) than the English set (21%), even though both sets contain many correct responses overall, demonstrating how directional asymmetries in AI translation are deeply rooted in the morphological and lexical typologies of the source and target languages.

4.3 Why Gemini Outperformed Copilot

The performance patterns observed in the dataset show that GG consistently outperformed MC not only in producing identical equivalents but also in supplying domain-appropriate terminology, resolving polysemy, and recognizing culturally embedded meanings. GG's ability to generate multiple correct equivalents, identify specialized uses in mathematics, engineering, medicine, textiles, and commercial domains, and interpret social or colloquial expressions indicates that its training data is both broader and more diverse. GG demonstrates domain sensitivity, idiomatic awareness, cultural knowledge, semantic shift ability, and multi-equivalent output. MC relies on form matching, literal mapping, orthographic merging, and lacks domain activation. The table reveals structural differences between the two models that go beyond accuracy percentages. By contrast, MC tended to produce literal translations, single-option equivalents, and context-neutral renderings, which lowered its identical-match accuracy even when the meaning was technically correct.

Developed within Google's ecosystem, GG draws on extremely large and diverse English corpora, including extensive web-scale text and high-frequency morphological patterns. These resources contain dense representations of the derivational structures that characterize attenuated metaphors, allowing GG to retrieve standardized equivalents with high accuracy and minimal lexical

variation. In contrast, MC is trained within Microsoft's ecosystem, which relies on licensed data, publicly available text, and human-generated material. While this data is exceptionally strong in technical, scientific, and academic domains—explaining

Further evidence supporting this explanation comes from GG's superior ability to decode distorted words in political images (Al-Jarf, 2026) and interpret calligraphic or stylized text (Al-Jarf, 2025c), tasks that require exposure to noisy, multimodal, and visually complex data. Google's training ecosystem includes massive amounts of such real-world visual-linguistic content, giving GG stronger robustness in OCR-like tasks and non-standard text forms. MC's training data is not optimized to the same extent for these multimodal distortions, which contributes to the performance gap observed in the present study.

A close examination of the error patterns reveals a fundamental asymmetry in MC 's linguistic capabilities: its English language processing and formal terminological repository are noticeably stronger than its Arabic comprehension. When analyzing the faulty responses in the English-to-Arabic direction, it becomes evident that MC 's errors are predominantly concentrated in specialized, culturally embedded, or commercial terminology. When faced with these challenging items, MC's default fallback strategy is strict literal translation. In almost all of its incorrect English-to-Arabic responses, the Arabic output generated by MC was a rigid, word-for-word translation that entirely missed the pragmatic or domain-specific target meaning.

By contrast, the performance patterns observed in the dataset reveal that GG consistently outperformed MC not only in producing identical correct equivalents but also in supplying domain-appropriate terminology, resolving polysemy, and recognizing culturally embedded meanings. GG demonstrated an advanced capacity to generate multiple correct equivalents, identify specialized uses across mathematics, engineering, medicine, textiles, and commercial domains, and accurately interpret social or colloquial expressions. This stark contrast highlights that GG possesses deep domain sensitivity, idiomatic awareness, cultural knowledge, and semantic shift ability. Conversely, MC relies heavily on surface-level form matching, literal mapping, and orthographic merging, often failing to activate the necessary contextual domain. These operational disparities reveal structural differences between the two models that go far beyond basic accuracy percentages.

This divergence in performance can be directly attributed to the distinct ecosystems and training corpora underlying each model. Developed within Google's ecosystem, GG draws on extremely large, diverse, web-scale corpora that encompass a massive spectrum of colloquialisms, regional dialects, market language, and high-frequency morphological patterns. These resources contain dense representations of the derivational structures that characterize attenuated metaphors, allowing GG to retrieve standardized equivalents with high accuracy and minimal lexical variation.

In contrast, MC is trained within Microsoft's ecosystem, which relies more heavily on licensed data, formal academic texts, publicly available official documents, and structured human-generated material. While this highly formal data pipeline makes MC exceptionally strong in strictly technical, scientific, and academic domains—explaining its high performance in previous studies on sleep terminology, chemical compounds, and formal medical vocabulary—it leaves the model less equipped to handle the specific morphological patterns, conversational nuances, and colloquialisms that define many half-expressions. Consequently, when encountering non-academic specialized terms, MC lacks the contextual breadth to infer the meaning, forcing it to default to literal translation or produce varied, less precise synonyms.

Further evidence supporting this ecosystem-based explanation comes from previous research evaluating AI performance on non-standard inputs. GG has demonstrated a superior ability to decode distorted words in political images (Al-Jarf, 2026) and interpret calligraphic or stylized Arabic text (Al-Jarf, 2025c)—tasks that require extensive exposure to noisy, multimodal, and culturally complex real-world data. Google's training ecosystem includes massive amounts of such diverse visual-linguistic content, granting GG stronger robustness in handling non-standard text forms and cultural nuances. MC's training data, being heavily optimized for formal and structured text, is simply not calibrated to the same extent for these socio-linguistic and multimodal variations, which perfectly aligns with the performance gap observed in this study.

Finally, GG's underlying attention mechanisms are structured to maintain deep contextual coherence over longer spans. In translation tasks involving dense technical domains (like medical, mathematical, or engineering half-terms), maintaining local-global consistency is vital. MC occasionally loses micro-context when parsing isolated or tightly packed multi-word technical units, leading to structural flattening or literal default outputs.

4.4 Comparison with Prior Studies

The findings of the present study differ substantially from most prior research on AI translation of metaphors, idioms, proverbs, and culturally loaded expressions. Earlier studies consistently reported that AI systems struggle with figurative language because metaphors are culturally embedded, semantically opaque, and often lack direct cross-linguistic equivalents. Studies on idioms and proverbs—such as Mughal et al. (2024), Shalevska et al. (2025), Raji Ellams (2025), and Al Salem et al. (2026)—showed that AI

frequently misinterprets nonliteral meaning, produces literal translations, or fails to capture cultural nuance. These results contrast sharply with the current study, in which MC and GG achieved high accuracy (80%) in translating Arabic half-expressions and moderate accuracy (56%) in translating English half-expressions. The difference arises because the expressions in this dataset are technical, literal, and semantically transparent, unlike idioms and proverbs that require cultural grounding.

Similarly, studies on literary metaphors—such as Karakanta, Nas & Dorst (2025), Fang & Wang (2025), Iskakbayeva et al. (2025), and Van Heerden & Bas (2024)—reported persistent AI difficulties with emotional depth, metaphorical complexity, and stylistic nuance. These findings diverge from the present results, where both models produced a high proportion of identical correct equivalents (61%) in the Arabic–English direction. The contrast highlights that AI performs significantly better when the input expressions have stable, standardized English morphological equivalents such as *semi-*, *hemi-*, *mid-*, and *middle*, rather than metaphorical or stylistic ambiguity.

Research on cultural metaphors—such as Azizov (2024), Vance Kaelen (2026), and Assiri (2026)—also differs from the current study. These works emphasize AI's tendency to misinterpret symbolic cultural meaning or produce culturally inappropriate equivalents. In contrast, the present dataset contains no culturally embedded metaphors; instead, the expressions denote quantitative, spatial, or temporal "half", which AI systems can map reliably to English morphological patterns. Thus, while prior studies highlight cultural misalignment, the current study demonstrates morphological alignment.

Studies in specialized domains, including tourism discourse (Spinzi, 2025), and religious metaphors (Farshi & Zare, 2025), similarly report challenges due to conceptual metaphor structures. These findings contrast with the present study, where the models' errors were not conceptual but lexical, arising from the non-standardized nature of Arabic terminology and the availability of multiple correct equivalents. This aligns with earlier observations that "equivalents are synonyms, a descriptive word or phrase," and that regional variation naturally occurs across Arabic.

Finally, computational metaphor-processing studies—such as Ge, Mao & Cambria (2023), Wang et al. (2024), and Bassenge (2026)—highlight the inherent difficulty of metaphor detection and interpretation. These challenges are largely absent in the current dataset because the expressions are non-metaphorical. Instead, the main difficulty lies in the English–Arabic direction, where Arabic offers multiple valid lexical choices. This finding aligns with Yaseen (2013) and other Arabic terminology studies showing that Arabic often provides several acceptable equivalents for a single English term.

In summary, prior studies consistently show that AI struggles with metaphorical, cultural, and idiomatic expressions, whereas the present study demonstrates that AI performs strongly when translating literal, technical, and morphologically regular expressions. The current results therefore contrast with metaphor-translation research but align with studies on terminological variation in Arabic, particularly those highlighting the lack of standardization and the richness of Arabic lexical options.

4.5 Comparison with the Author's Studies on MC translations

The results of the present study show that MC performed substantially better on half-expressions than on many of the linguistic domains examined in the author's previous research. In the current dataset, MC produced 80% correct equivalents for Arabic half-expressions and 65% correct equivalents for English half-expressions. These figures are noticeably higher than MC's accuracy in earlier studies involving contrastive, emphatic, and negative expressions, where MC achieved 65% correctness. The improvement suggests that half-expressions—being literal, structurally transparent, and morphologically regular—align more closely with MC's strengths in lexical matching and pattern recognition.

Across the author's prior studies, MC's accuracy varied widely depending on the linguistic domain. High performance was observed in areas with fixed, standardized terminology, such as English sleep terms (91%), English formulaic expressions (79%), common chemical compound names (72%), and medical terminology (68.6%). These domains rely on stable lexical equivalents that appear frequently in training corpora, enabling MC to retrieve them reliably. The current study aligns with these high-accuracy domains because half-expressions also map onto predictable English morphological patterns such as *semi-*, *hemi-*, *mid-*, and *middle*.

In contrast, MC's performance dropped sharply in domains requiring figurative interpretation, cultural knowledge, or contextual inference. These include expressions of impossibility (52%), Arabic folk medical terms with *abu* and *umm* (46%), metaphorical Arabic grammatical terms (43%), Gaza–Israel terminology (29%), and Arabic *abu-brand* names (0%). Similarly, MC struggled with metonymic *abu/umm* animal and plant names, producing high rates of faulty equivalents, literal translations, and misreadings of metonymic names as personal names. These results contrast with the current study, where MC produced a high proportion of identical correct equivalents (61%) in the Arabic–English direction, indicating that half-expressions do not require cultural grounding or metaphorical decoding.

MC's performance in decoding encrypted Arabic on social media also showed variability: 56% correct in general encrypted text, 16% partial and 27% faulty in political content, and mixed accuracy in COVID-19 discourse. These findings differ from the current study, where MC's errors were primarily lexical rather than contextual or pragmatic. The half-expressions dataset does not involve political, cultural, or metaphorical ambiguity, which explains the higher accuracy.

Finally, the variation across studies reflects both domain-specific linguistic demands and temporal improvements in MC's capabilities. Over the two-year period during which the author conducted these investigations, MC's performance improved due to system updates, expanded training data, and iterative refinement. Sustained interaction with the author—through repeated corrections, discussions, and prompting—functioned as informal user-driven fine-tuning, contributing to more stable and contextually appropriate outputs. The current study benefits from these improvements, as well as from the inherently literal and morphologically regular nature of half-expressions.

4.6 Impact of Blind Testing, Exhaustive Equivalent Prompting, and Isolated Presentation

An important methodological dimension of this study is that the AI models were tested under blind conditions**:** the prompts instructed the systems to provide as many valid equivalents as possible for each expression without prior classification, contextual labeling, or hints regarding whether the items were folk, general, or specialized. Furthermore, the half-expressions were presented to the models entirely in isolation, stripped of any sentential, textual, or pragmatic context.

This unguided, open-ended prompt design combined with isolated presentation ensured that the models' performance relied entirely on their internal semantic representations and cross-linguistic retrieval mechanisms rather than artificial categorization constraints or contextual scaffolding. Without surrounding text to narrow down meaning, the models were forced to rely solely on static token-level associations. This explains the lower identical-match accuracy in the English-to-Arabic direction, as the lack of context forced both models—particularly Arabic's lexical richness—to generate diverse stylistic variants and synonyms (نصف, شبه, منتصف, وسط). Consequently, the observed variations reflect the true architectural and corpus-based limits of each system when processing mixed, real-world data streams entirely out of context.

4.7 Limitations of Copilot

MC's performance reveals several structural limitations. First, the model relies heavily on literal form-matching, often mapping half to نصف even when the English expression requires semantic shift or idiomatic rendering. Second, MC shows weak domain activation, producing literal translations in specialized fields where GG retrieves precise technical equivalents. Third, MC frequently merges compound terms orthographically (e.g., halfcocked, half coupling, half-hardy), indicating limited sensitivity to English morphological conventions. Fourth, the model struggles with culturally embedded expressions, both Arabic and English, including folk terms, slang, and idioms such as half seas over, half cut, half proof. Fifth, MC lacks the ability to provide multiple equivalents or contextual explanations, offering single literal outputs even when the expression is polysemous or context-dependent. These limitations suggest that MC's Arabic coverage, idiomatic knowledge, and cultural grounding are significantly weaker than its English lexical mapping capabilities.

4.8 Implications for AI Translation

The findings reveal that AI translation performance is strongly shaped by domain sensitivity, cultural knowledge, and the ability to perform semantic shifts rather than literal form-matching. GG's consistent activation of domain-specific knowledge across medicine, engineering, music, navigation, sports, and folk usage demonstrates that effective AI translation requires rich, diversified training corpora and robust domain-linking mechanisms. The study also shows that idiomatic and culturally embedded expressions—particularly attenuated metaphors such as half-expressions, folk terms, and market-based terminology—cannot be translated through morphology or surface structure alone. This highlights the need for AI systems to integrate cultural lexicons, idiom databases, and pragmatic inference models. The asymmetry between Arabic-to-English and English-to-Arabic translation further indicates that AI models perform best when expressions have stable one-to-one equivalents, and struggle when translation requires cultural interpretation, derivational awareness, or semantic restructuring.

5. Recommendations

5.1 Recommendations for Developers

Based on the performance patterns and limitations observed in the present study, several key recommendations can be offered to improve the accuracy and reliability of AI-generated equivalents for Arabic and English half-expressions:

- AI developers should broaden Arabic training datasets to include colloquial usage, folk expressions (such as regional derivations like نص انصيص), marketplace terminology, and culturally embedded metaphors, moving beyond a primary reliance on Modern Standard Arabic.

- Models should incorporate structural awareness of Arabic root systems (such as ف-ص-ن), enabling them to recognize that a single root generates multiple morphological forms whose functions cannot be mapped directly onto English morphology, paraphrastic constructions, or zero-derived verbs.
- Training datasets should include robust representations of literal, morphologically regular expressions, such as half-expressions, zero-expressions, and contrastive structures. The high accuracy observed in this study confirms that AI systems heavily benefit from exposure to predictable morphological patterns.
- Developers must integrate advanced semantic disambiguation mechanisms to address challenges in the English–Arabic direction, where Arabic's lexical richness often yields multiple valid synonyms (نصف, شبه, منتصف, وسط). Enhanced disambiguation would help models select the most contextually appropriate equivalent rather than a merely plausible variant.
- AI systems should incorporate dedicated idiom and metaphor databases, allowing models to automatically distinguish between literal mathematical fractions, specialized commercial nomenclatures, and attenuated metaphorical usage.
- Across prior studies, AI performance improved noticeably over time due to iterative prompting, correction, and sustained interaction. Incorporating structured user feedback loops and adaptive learning pipelines into model updates would support greater stability and consistency across complex linguistic domains.
- Developers should improve the orthographic processing of English compounds to prevent the erroneous merging of terms, thereby enhancing the model's ability to detect whether half functions as a prefix, a modifier, an idiom, or a technical label.

Together, these recommendations highlight the critical need for improved modeling of Arabic lexical variation, expanded training data for predictable morphological structures, enhanced figurative-literal discrimination, stronger semantic disambiguation, and systematic user-guided refinement. Implementing these strategies will strengthen AI translation performance not only for half-expressions, but across broader Arabic–English translation tasks.

5.2 Recommendations for Students on Using AI in Translating Half-Expressions

When utilizing artificial intelligence tools to translate metaphorical expressions in general and half-expressions in particular, students must adopt a critical, analytical approach rather than treating AI outputs as infallible authorities. Based on the performance patterns and limitations observed in this study, students are advised to follow the guidelines below:

- Students should use AI tools to quickly retrieve standardized morphological patterns (such as semi-, hemi-, mid-, or middle) and generate initial drafts, but they must cross-reference outputs against reliable bilingual and monolingual dictionaries, specialized glossaries, and domain-specific resources to ensure complete accuracy.
- Because AI models—particularly MC—frequently default to literal, form-based translations when encountering specialized, commercial, or regional half-expressions, students must verify whether an expression functions as a specialized nomenclature or a fixed lexical unit rather than a simple mathematical fraction or word-for-word combination.
- When translating from English to Arabic, students should expect AI models to generate multiple valid synonyms or stylistic variants (such as choosing between نصف, شبه, منتصف, or وسط). Students must exercise contextual judgment to select the precise equivalent suited to the specific domain rather than accepting the first AI-generated option.
- Arabic half-expressions and related verbs (such as شطر, قسمه, فلقه) should not be translated imitatively. Because Arabic is a highly derivational language generating multiple forms from the root ف-ص-ن (such as نصف, نصفين, مناصفة, تناصف, ناصف), certain structures—like reciprocal forms (تناصف)—have no zero-derived English verbs. Students must recognize that English expresses these meanings through distinct lexicalized forms (semi-, hemi-, demi-, di-, bi-), paraphrastic constructions, or unrelated verbs (split, divide, truncate) rather than direct morphological mirroring.
- AI models struggle significantly with culturally embedded folk expressions, oral folklore, and non-standard derivations (such as نص انصيص). Students should not rely on AI to decode expressions rooted in regional culture or folklore, as these require human cultural grounding rather than algorithmic mapping.
- Students should always supply the domain of the half-expression (e.g., dentistry, textiles, navigation, electronics) and request definitions alongside translations to distinguish literal, idiomatic, and technical meanings, as domain activation significantly improves accuracy.
- Students should ask AI to generate multiple translation options for comparison, request reverse-translation checks to ensure semantic stability, and use the tool to compare half-expressions across dialects, registers, and geographic variants (such as British versus American English or formal versus informal tones).
- Students should build a habit of consulting multiple reliable resources to verify AI output. Exposure to standardized terminology, combined with regular practice, peer discussion, and instructor feedback, enhances their ability to select precise equivalents and handle structurally complex expressions with confidence.
- By continuously evaluating whether AI has captured the correct domain context or fallen into a literal translation trap, students can leverage AI efficiency while safeguarding accuracy, contextual precision, and cultural authenticity.

6. Conclusion

The present study examined how MC and GG translate Arabic and English half-expressions. The results revealed that both models performed strongly, with GG producing 92% correct equivalents for Arabic half-expressions (compared to 80% for MC) and 80% for GG (compared to 65% for MC) for English half-expressions. GG outperformed MC in identical-match accuracy, particularly in the Arabic–English direction, reflecting its stronger alignment with standardized English morphological patterns such as *semi-*, *hemi-*, *mid-*, and *middle*. The study also showed that both models avoided incorrect prefixes such as *bi-* and *di-*, and occasionally used *demi-* in fixed lexical items.

These findings contrast with the author's earlier studies, where MC's accuracy fluctuated widely. High performance was observed in standardized technical terminology, while culturally embedded, metaphorical, or metonymic expressions produced substantially lower accuracy. The current study demonstrates that AI systems perform best when the translation unit is literal, predictable, and morphologically regular, whereas performance declines when expressions require cultural knowledge, figurative interpretation, or contextual inference. The results therefore highlight the importance of linguistic domain, expression type, and morphological transparency in shaping AI translation accuracy.

One limitation of the current study is that it examined only two AI systems, with GG evaluated on fixed and formulaic expressions for the first time, unlike MC, which was the primary focus of the author's prior studies on metaphorical expressions. Therefore, future research should broaden its scope to examine the performance of additional models such as Claude, DeepSeek, and Qwen.

Furthermore, because the author has conducted numerous studies on AI translation using MC in particular, it remains unclear whether the high accuracy observed for half-expressions stems from MC's architectural improvements or user-driven refinement resulting from the model's memory feature. Throughout prior studies, MC's accuracy improved noticeably over time, raising the question of whether this progress reflects genuine developer updates or localized adaptation driven by sustained interaction with the author, repeated corrections, and contextual familiarity. Comparing outputs under controlled conditions—specifically with system memory enabled versus disabled—would help determine whether high performance stems from intrinsic model advancements or user-specific tuning. Controlled experiments across diverse linguistic domains (literal, metaphorical, technical, and cultural) would provide valuable insights into how AI systems learn, adapt, and stabilize over time.

Ultimately, these findings underscore the need for continued, domain-sensitive evaluation of AI translation systems. As the present study demonstrates, performance is not uniform across linguistic categories; rather, it is shaped by the structural, morphological, and cultural characteristics of the expressions being translated. Understanding these dynamics will support more accurate, reliable, and context-appropriate AI translation in future applications.

Conflicts of Interest: The author declares no conflict of interest.
ORCID ID: https**:**//orcid.org/0000-0002-6255-1305

References

[1] Al-Jarf, R. (2025a). AI translation of full-text Arabic research articles: The case of educational polysemes. *Journal of Computer Science and Technology Studies*, *7*(1), 311-325. DOI: 10.32996/jcsts.2025.7.1.23

[2] Al-Jarf, R. (2025b). AI translation of the Gaza-Israel war terminology. *International Journal of Linguistics, Literature and Translation*, *8*(2), 139-152. DOI: 10.32996/ijllt.2025.8.2.17

[3] Al-Jarf, R. (2025c). Calligraphic text recognition by Gemini, Ernie ViLG and Google Translate: A comparative study of Arabic, Japanese and Chinese. *Journal of Computer Science and Technology Studies*, *7*(12), 474-494. DOI: 10.32996/jcsts.2025.7.12.54

[4] Al-Jarf, R. (2025d). *Can Artificial Intelligence (AI) translate Arabic abu-brand names with different prompts*. *Journal of Computer Science and Technology Studies*, *7*(9), 768-779. DOI: 10.32996/jcsts.2025.4.1.91.

[5] Al-Jarf, R. (2025e). *Can ESP students use Artificial Intelligence for translating common names of chemical compounds*. 6th INNO-LSP International Conference titled: Innovative Trends in Teaching Languages For Specific Purposes in the Context of the Current Social Instability. Ovidius University of Constanța, Romania. September 18- 19, 2025. https://www.researchgate.net/publication/398423410. Google Scholar

[6] Al-Jarf, R. (2025f). Copilot's English translation of contrastive emphatic negation in Arabic discourse: An analytical study. *International Journal of Linguistics, Literature and Translation*, *8*(12), 214-230. DOI: 10.32996/ijllt.2025.8.12.24

[7] Al-Jarf, R. (2025g). Copilot vs DeepSeek's translation of denotative and metonymic abu- and umm- animal and plant folk names in Arabic. *Journal of Computer Science and Technology Studies*, *7*(10), 367-385. DOI: 10.32996/jcsts.2025.7.10.39

[8] Al-Jarf, R. (2025h). 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

[9] Al-Jarf, R. (2025i). Google Translate then and now: Translations from five languages into English and Arabic (2012–2025). *Journal of Computer Science and Technology Studies*, *7*(12), 413-427. DOI: 10.32996/jcsts.2025.7.12.50

[10] Al-Jarf, R. (2025j). Human vs AI translation of common names of chemical compounds: A comparative study. Frontiers in Computer Science and Artificial Intelligence, 4(4), 11-24. DOI: 10.32996/fcsai.2025.4.4.2

[11] Al-Jarf, R. (2025k). Translation of Arabic folk medical terms with om and abu by AI: A comparison of Microsoft Copilot and DeepSeek. *Journal of Medical and Health Studies*, *6*(4), 45-58. DOI: 10.32996/jmhs.2025.6.4.8. Google Scholar

[12] Al-Jarf, R. (2025l). 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. DOI: 10.32996/jcsts.2025.7.8.33

[13] Al-Jarf, R. (2025m). Translation of English and Arabic "sleep" terms and formulaic expressions by Artificial Intelligence: A comparison of Copilot and DeepSeek. *International Journal of Linguistics, Literature and Translation*, *8*(11), 95-108. DOI: 10.32996/ijllt.2025.8.11.10.

[14] Al-Jarf, R. (2025n). *Translation of "sleep" terms & formulaic expressions by Artificial Intelligence*. VI International Conference "Metaphor in Education". Faculty of Modern Foreign Languages and Literatures. Perm State University, Russia. October 15-17, 2025. https://www.researchgate.net/publication/398665426. Google Scholar

[15] Al-Jarf, R. (2025o). Translation of zero-expressions by Microsoft Copilot and Google Translate. *Journal of Computer Science and Technology Studies (JCSTS)*, *7*(2), 203-216. DOI: 10.32996/jcsts.2025.7.2.20

[16] Al-Jarf, R. (2024a). *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, pp. 1–16, 2024. Springer DOI: 10.1007/978-3-031-71429-0_17.

[17] Al-Jarf, R. (2024b). Trans*lation of medical terms by AI: A comparative linguistic study of Microsoft Copilot and Google Translate*. I2COMSAPP'2024 Conference, 23-25 April. Nouakchott, Mauritania. https://www.researchgate.net/publication/387108997 Google Scholar

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

[19] 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. https://www.researchgate.net/profile/Reima-Al-Jarf/publication/346471862. Google Scholar

[20] Ahmed, O., et al. (2025). Translating Arabic Literary Metaphors: A Comparative Evaluation of ChatGPT-4.0 and Google Translate 2024. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, *13*(2), 3177-3200.

[21] Al Salem, M., et al. (2026). AI-powered translation of Arabic idioms into English. *Training, Language and Culture*, *10*(1), 8-21.

[22] Alazzam, T., et al. (2025). Exploring AI's Capability in Translating English Metaphors Into Arabic. *Theory & Practice in Language Studies (TPLS)*, *15*(7).

[23] Assiri, B. (2026). Lost in Translation: AI Recognition and Misinterpretation of Culturally Embedded Symbolic Language. In *AI Bots in Applied Linguistics: Practice, Policy, and Creativity* (pp. 315-346). IGI Global Scientific Publishing.

[24] Azizov, D. (2024). From Idioms to Algorithms: Translating Culture-Specific Expressions in AI Systems. *Iconic Research And Engineering Journals*, *7*(10), 543-551.

[25] Bassenge, M. (2026). Innovation or Imitation? Assessing Novelty in AI-Generated Metaphors through Automated Conceptual Metaphor Analysis with MetaNet, Word Embeddings, and Conceptual Mapping Techniques. Master's thesis. Universitäts- und Landesbibliothek Darmstadt, Germany.

[26] Brala Vukanović, M. (2025). Translating (metaphors) in the age of AI: Opportunities, challenges, and implications for the EFL classroom. *ELOPE, 22*(1), 185-201. DOI : 10.4312/elope.22.1.185-201.

[27] Darwish, N., et al. (2025). A reception study of AI-translated idioms and proverbs between Arabic and English. *Humanities*, *6*(3).

[28] Erokhin, V. (2026). Assessment of English-Russian MT systems: translation of conceptual metaphors. Master's Thesis. Vytautas Magnus University, Lithuania.

[29] Fang, H. & Wang, B. (2025). Assessing Machine Translation of Emotional Depth, Metaphorical Complexity, and Cultural Nuances in Literary Texts—A Case Study of Dream of the Red Chamber. *International Journal of Chinese and English Translation & Interpreting*.

[30] Farshi, V., & Zare, S. (2025). A critical evaluation of the translation of conceptual metaphors of light and darkness in the Qur'an by AI models in light of Lakoff and Johnson's theory and newmark's strategies. *Translation Researches in the Arabic Language And Literature*, *15*(33), 261-295.

[31] Ge, M., et al. (2023). A survey on computational metaphor processing techniques: From identification, interpretation, generation to application. *Artificial Intelligence Review*, *56*(Suppl 2), 1829-1895.

[32] Guo, L., & Amirdabbaghian, A. (2025). Exploring the translation of idioms by AI in Chinese internet literature. In *Role of AI in Translation and Interpretation* (pp. 293-326). IGI Global Scientific Publishing.

[33] He, Y., & Jiang, C. (2024). Cultural perspectives on the translation system of political text metaphors using Artificial Intelligence research. *Computer-Aided Design & Applications*, *21*, 207-219.

[34] Iskakbayeva, A., et al. (2025). Comparative analysis of literary translation using Artificial Intelligence: An experiment based on EA Poe's short story "The Tell-Tale Heart". *Dibon Journal of Languages*, *1*(3), 304-313.

[35] Karakanta, A., et al. (2025). Metaphors in Literary Machine Translation: Close but no cigar?. In *Proceedings of Machine Translation Summit XX: Volume 1* (pp. 276-286).

[36] Kūlis, M. (2026). Lost in Translation: Artificial Intelligence and the Burden of Bad Metaphors. In Vincent C. Müller, Leonard Dung, Guido Löhr & Aliya Rumana, Philosophy of Artificial Intelligence: The State of the Art. Berlin: Springer Nature.

[37] Mughal, U., et al. (2024). The intersection of linguistics and artificial intelligence: a corpus-based study of idiom translation. *Journal of Applied Linguistics and TESOL (JALT)*, *7*(4), 1453-1460.

[38] Nafiza, S. & Erwina, E. (2025). Translation of metaphors in ex machina from English to Indonesian. *International Journal Linguistics of Sumatra and Malay*, *4*(2), 62-69.

[39] Ordóñez, I. (2026). Metaphor Translation Revisited in the Age of AI: A Case Study with English Financial and Economic Reports Translated into Spanish. *Translating Finance and Economics: Theory, Practice, Training and Technology*, 165.

[40] Raji-Ellams, Z. B. Cross-linguistic interpretation of metaphor in Yoruba proverbs: a pragma-cognitive ai-assisted study. Ilorin Journal of Translation Studies by Institute of Translation Arts, Unilorin. 1-26.

[41] Serón Ordóñez, I. (2026). Metaphor Translation Revisited in the Age of AI: A Case Study with English Financial and Economic Reports Translated into Spanish and French. In *Translating Finance and Economics: Theory, Practice, Training and Technology* (pp. 165-201). Singapore: Springer Nature Singapore.

[42] Shalevska, E., et al. (2025). AI in nonliteral language translation: Translating Macedonian proverbs and idioms. *International journal of Education Teacher*, *29*, 60-66.

[43] Skrynnikova, I. (2024). Interpreting metaphorical language: a challenge to artificial intelligence. *Вестник Волгоградского государственного университета. Серия 2: Языкознание*, *23*(5), 99-107.

[44] Spinzi, C. (2025). Translating the intangible: The role of Artificial Intelligence in the translation of metaphors in tourism discourse. *International Journal of Language Studies*, *19*(2).

[45] van Heerden, I. & Bas, A. (2024). A perspective on literary metaphor in the context of generative AI. *arXiv preprint arXiv:2409.01053*.

[46] Vance-Kaelen, Z. (2026). Algorithmic Translation of Cultural Metaphors: A Conceptual Synthesis of Cognitive Semiotics and AI Pragmatics. Guided Teaching and Learning Research, 2(1), 1-5.

[47] Wang, S., et al. (2024). MMTE: Corpus and metrics for evaluating machine translation quality of metaphorical language. In *Proceedings of the 2024 conference on empirical methods in natural language processing* (pp. 11343-11358).

[48] Wang, Z. & Chai, J. (2024). On metaphor translation into English based on artificial intelligence. *Procedia Computer Science*, *247*, 1359-1365.

[49] Yu, Y. & Shao, L. (2023). A cognitive stylistic approach to the translation of spatial metaphors in Alai's Chen'ai luoding. *Asia Pacific Translation and Intercultural Studies*, *10*(1), 20-33.