

## **The Efficacy of AI-Powered Translation of English News Reports on the Gaza Conflict into Arabic**

**Murtadha Hasan Abed Ali**

**Prof. Dr. Mahmood Ibrahim Hamdan**

**University of Tikrit/ College of Arts/ Department of Translation**

**[MH230075prt@st.tu.edu.iq](mailto:MH230075prt@st.tu.edu.iq)**

### **Abstract**

The study investigates the translation gaps produced by AI models when rendering politically sensitive English news reports into Arabic, using the Gaza conflict as a case study. While GPTs demonstrate remarkable efficiency, their outputs sometimes lack the cultural nuance and emotional depth required for accurate translations of contentious geopolitical discourse. This research aims to analyse the differences between AI-generated and human translations in politically sensitive contexts and examine variations in context, nuances, and language between GPT and human outputs. The study hypothesises that GPT-generated translations will exhibit significant gaps in accuracy and nuance compared to human translations. It adopts Descriptive Translation Studies (DTS) by Gideon Toury and Halliday's Systemic Functional Linguistics (SFL) as its theoretical models. The study relies on GPT-40-mini for AI translations, documenting the translation date due to its continuous evolution. A comparative analysis of each GPT and human translation pair focuses on semantic accuracy, handling politically sensitive terms, cultural appropriateness, and detecting translation gaps and errors. This study is concerned that GPT algorithms, trained on large databases, often reflect biases related to culturally sensitive situations. This study concludes that GPT-generated translations often lack the cultural context, nuance, and emotional depth provided by human translators, leading to accuracy gaps. Furthermore, complete removal of translation gaps in politically sensitive content will be impossible in the near future, despite targeted algorithmic improvements. This research contributes to improving AI-driven translation accuracy, particularly in vital contexts like conflict reporting, and underscores the need for refined AI models to ensure reliable news dissemination across English and Arabic audiences.

Keywords: (AI-Powered Translation, GPT tool, political discourse, translation gaps).

## فعالية الترجمة المدعومة بالذكاء الاصطناعي في نقل التقارير الإخبارية الإنجليزية حول صراع غزة إلى العربية

الباحث مرتضى حسن عبد علي

أ.د. محمود إبراهيم حمدان

جامعة تكريت/ كلية الآداب/ قسم الترجمة

[MH230075prt@st.tu.edu.iq](mailto:MH230075prt@st.tu.edu.iq)

### الملخص

تبحث الدراسة الفجوات الترجمة التي تنتجها نماذج الذكاء الاصطناعي عند ترجمة التقارير الإخبارية الإنجليزية المشحونة سياسياً إلى اللغة العربية، باستخدام صراع غزة كدراسة حالة. بالرغم من أن نماذج (جات جي بي تي) تُظهر كفاءة ملحوظة، إلا أن مخرجاتها تفتقر أحياناً إلى الدقة الثقافية والعمق العاطفي المطلوبين لترجمة الخطاب الجيوسياسي الخلفي بدقة. وتهدف هذه الدراسة إلى تحليل الفروق بين الترجمات التي ينتجها الذكاء الاصطناعي والترجمات البشرية في السياقات السياسية الحساسة، إضافة إلى الاختلافات السياقية والفروق الدلالية واللغة بين مخرجات (شات جي بي تي). تفترض الدراسة أن الترجمات المؤلدة بواسطة (جات جي بي تي) تُظهر فجوات كبيرة في الدقة والفروق الدلالية مقارنة بالترجمات البشرية. تعتمد الدراسة على إطارين نظريين رئيسيين هما؛ دراسات الترجمة الوصفية (DTS) لجيدون توري، واللسانيات الوظيفية النظامية (SFL) لهاليداي. استخدمت الدراسة نموذج-GPT 4o-mini للحصول على الترجمات الآلية، مع توثيق تاريخ الترجمة نظراً للتطور المستمر للنموذج. يُعتمد إجراء تحليل مقارن لكل زوج من الترجمات (المنتجة بواسطة GPT والبشرية)، مع التركيز على الدقة الدلالية، والتعامل مع المصطلحات المشحونة سياسياً، والملاءمة الثقافية، ورصد الفجوات والأخطاء الترجمة. وتشير الدراسة قفلاً من أن خوارزميات GPT ، المُدرّبة على قواعد بيانات ضخمة، تعكس في كثير من الأحيان تحيزات مرتبطة بسياقات حساسة ثقافياً. تخلص الدراسة إلى أن الترجمات المؤلدة بواسطة GPT تفتقر في كثير من الأحيان إلى السياق الثقافي والفروق الدلالية والعمق العاطفي الذي توفره الترجمات البشرية، مما يؤدي إلى فجوات في الدقة. علاوة على ذلك، فإن إزالة الفجوات الترجمة تماماً في المحتوى السياسي الحساس ستبقى مستحيلة في المستقبل القريب، رغم التحسينات الخوارزمية المستمرة. وتوصي هذه الدراسة في تحسين دقة الترجمة المدعومة بالذكاء الاصطناعي، خاصة

في السياقات الحيوية مثل تغطية النزاعات، كما تؤكد الحاجة إلى نماذج ذكاء اصطناعي أكثر تطوراً لضمان نشر إخباري موثوق بين الجمهور الناطق بالإنجليزية والعربية.

الكلمات المفتاحية: (الترجمة المدعومة بالذكاء الاصطناعي، أداة (جي بي تي)، الخطاب السياسي، فجوات الترجمة).

## Introduction

### 1.1 Statement of the Problem

The rapid advancement of artificial intelligence (AI) in machine translation improved efficiency in language communication, yet there are still significant gaps in accurately rendering politically sensitive content, especially in high-risk contexts like conflict reporting. While AI-powered tools such as GPT demonstrate proficiency in general translation tasks, their ability to handle precise, culturally embedded, and ideologically charged language—especially in translating news from English to Arabic—remains understudied. This study investigates the translation variations produced by English-to-Arabic translation tools when processing English news reports on the Gaza conflict, a highly sensitive geopolitical topic that requires precision in terminology, tone, and ethical framing.

Current research highlights AI's limitations in contextual awareness, pragmatic accuracy, and bias mitigation, particularly in languages with complex socio-political connotations like Arabic. The study aims to bridge this gap by examining English source texts from authoritative broadcasters and institutions—including BBC Arabic, CNN, DW, Al Jazeera, official government statements, select newspapers, and United Nations (UN) documents—alongside their professionally translated Arabic versions. These human translations serve as a standard to assess the accuracy and reliability of GPT-generated translations.

By focusing on the Gaza conflict as a case study, this research underscores the risks of relying on only AI for translating contested narratives, where even minor inaccuracies can amplify misinformation or misrepresent stakeholder perspectives. The findings aim to contribute to translation studies by identifying systemic weaknesses in AI-powered news localization

while proposing strategies to minimize these gaps in politically charged discourse.

## 1.2 Aims of the Study

This study aims to the following:

1. Examine the differences in context, nuances, and language between translations provided by GPT and those performed by people.
2. Improve AI translation tools by giving suggestions on how to make AI translation algorithms better, especially when it comes to dealing with sensitive and complicated subjects.
3. Provide perspectives on the implications of the use of AI in translation, especially in politically charged environments, thereby contributing to the broader field of translation studies.

## 1.3 Questions of the study

1. In what ways might GPT's translations reflect inherent biases or neutrality challenges, particularly in the context of sensitive geopolitical discussions?
2. In which specific areas—such as fluency, terminological consistency, and contextual appropriateness—does GPT's translation performance fall short compared to human translators?
3. To what extent does GPT provide accurate and reliable translations of English news reports on the Gaza conflict into Arabic?

## 1.4 Hypotheses of the Study

This study hypothesizes that:

1. Targeted improvements in GPT translation algorithms can totally remove all translation errors, especially in politically sensitive topics.

2. GPT's algorithms are pre-trained on a huge database and are therefore biased towards culturally sensitive topics, specifically the conflict in Gaza.
3. The study hypothesizes that AI-powered translation of previously published online news yields more accurate results compared to unpublished data.

### **1.5 The Models Adopted**

In this study the researcher adopts two models: Descriptive Translation Studies (DTS) by Gideon Toury and Halliday's Systemic Functional Linguistics (SFL). Toury's (DTS) approach emphasizes the analysis of translations within their sociocultural contexts, focusing on how translations function in the target culture rather than adhering strictly to source texts. Halliday's (SFL) offers a framework for understanding language as a social semiotic system, highlighting how language choices reflect social contexts and communicative purposes. Together, DTS and SFL deepen understanding of translation and language use in cultural communication.

### **1.6 Procedures and Data Collection**

1. The data is selected from a range of English-language news reports on the Gaza conflict from official bilingual sources (e.g., BBC, Reuters, DW, CNN, and Al Jazeera) to ensure reliability. To guarantee translation reliability, only reports that include both the original English texts and their official human-translated Arabic equivalents were taken into examination.
2. The study relies on “GPT-40-mini—an AI tool developed by OpenAI”. This version considers that at the time of writing the study is considered one of the latest versions of GPT. The date of the GPT translation is documented because it is constantly evolving.

3. Regarding each news title, a triangular table is constructed that includes (a) the original English text, (b) the AI-produced Arabic translation, and (c) the human-produced Arabic translation.
4. Making a comparative analysis of each pair of GPT and human translations, focusing on semantic accuracy, handling politically sensitive terms, cultural appropriateness, and detecting translation gaps and errors in the GPT output.

### **1.7 Scope of the Study**

The scope of the study is limited to news related to the conflict in Gaza as a politically charged topic. The study relies on news produced by news institutions like CNN, DW, BBC, some international newspapers, and UN organisations ; these institutions produce their news in more than one language, especially English and Arabic. The Arabic translation is certified by these institutions as the standard for translation and compared to the GPT translation. A comparative analysis framework is adopted, where GPT translations are critically assessed against human translations in terms of semantic coherence, pragmatic appropriateness, and contextual fidelity.

### **1.8 Significance of the Study**

The study analytically examines the limitations of GPT in translating politically sensitive content, contributing to improving AI-driven translation accuracy, particularly in vital contexts like conflict reporting. Insights from this research can guide media organizations in integrating AI tools more cautiously, ensuring translations preserve intended meaning without ideological distortion. Findings highlight risks of AI-generated mistranslations in shaping public perception, advocating for ethical considerations in deploying automated tools for sensitive geopolitical discourse.

## 2 Literature Review

### 2.1 AI and Translation

The growing reliance on AI-powered machine translation tools is transforming the traditional methods of disseminating news through linguistic communication, particularly in war zones where accurate reporting is essential. The challenges persist in ensuring that the translation results uphold not only linguistic accuracy but also account for cultural and contextual variations, especially regarding sensitive topics such as the Palestinian issue. The war on Gaza is a highly controversial topic, as Western media often plays a role in shaping worldviews before translating them into Arab societies (Baker, 2018, p. 45). Due to the complexity of the vocabulary surrounding the Palestinian issue and the potential for ideas and concepts to be presented or interpreted in ways that align with a particular ideology, AI translations must balance subtle linguistic nuances with ethical issues. Studies show that while AI generally performs well at translation, it faces challenges in dealing with specialized language, hidden biases, and rhetorical differences that impact the recipient (Kenny, 2021, p. 78).

### 2.2 Cultural and Linguistic Distinctions in Arabic and English News

Translating news between Arabic and English is a complex task that involves not just linguistic skills but also an understanding of cultural contexts and ideological perspectives. To truly capture the accuracy and subtlety of the message, human translators are essential. While AI can certainly help speed up the process, it currently struggles with certain challenges, such as the variety of dialects, politically charged language, and the nuanced rhetoric often found in high-context discussions. To improve AI translation, future developments should focus on creating ethically sourced multilingual datasets and enhancing the ability of models to adapt to

different contexts. This approach will help reduce biases and improve the quality of cross-cultural news sharing.

### 2.3 Traditional vs AI-Powered Translation: A Theoretical Contrast

The fundamental distinction between traditional and AI translation lies in their underlying theoretical frameworks and operational methodologies. Traditional translation aligns with hermeneutic and phenomenological approaches that emphasize the subjective, interpretive nature of language (Gadamer, 1975, p. 162), requiring human translators to employ creative paraphrasing and cultural adaptation rather than literal substitution, particularly for idiomatic expressions and stylistic nuances (Williams, 2019, p. 78). This human-centric process is inherently multidisciplinary, involving iterative reviews and expert consultations to ensure fidelity to source content. In contrast, AI translation is rooted in computational linguistics and information theory, treating language as a system of statistical patterns (Manning & Schütze, 1999, p. 45). While AI systems demonstrate remarkable processing speed and scalability, their reliance on deep learning models often sacrifices contextual preservation and cultural adaptation (Garcia, 2021, p. 120).

Key limitations of AI translation include its inability to handle conceptual accuracy, terminological precision, and creative adaptation, particularly for under-represented

languages. Unlike human translators who can adapt texts to evolving linguistic trends and client specifications, AI systems depend on static algorithms that require frequent retraining (Johnson & Lee, 2022, p. 55). Ethical concerns further differentiate these approaches: human translators exercise discernment with sensitive content, while AI systems risk perpetuating pre-existing biases, such as in gender-neutral term translation (Müller, 2023, p. 212). The lack of socio-political awareness in AI systems

underscores the continued necessity of human oversight, particularly for specialized domains and historically significant texts. These comparative analyses reveal that while AI excels in efficiency, human translation remains superior in cultural, ethical, and adaptive dimensions.

## 2. 4 Human-AI Translation

The combination of artificial intelligence and language translation has revolutionized translation practices, improving translation accuracy and performance. With significant advances in AI tools, they have become a viable alternative to human translators as a translation assistant. This collaborative process between human translators and AI translators allows them to leverage these tools while preserving the accuracy of human translation.

As stated by O'Hagan, the combination of human translators and AI-powered translation, such as machine translation, makes the translation process highly flexible. AI translation can be used as a preliminary draft of human translation, and the results can be further refined to ensure cultural and contextual consistency. This process not only speeds up translation but also ultimately reduces the intellectual burden on the human translator, drawing attention to the most challenging aspects of the text (O'Hagan, 2016, pp. 130, 134). Additionally, during the translation process, the human translator refines some of the corrections to the machine translation results. These corrections continuously improve the machine translation in the future, along with the role of human feedback, thus aligning it more reliably with human linguistic sensibilities (Koehn, 2020, p. 215).

## 3 Methodology

The methodology used in this study is searching news headlines with details or news headlines without details. The news headlines related to the

war on Gaza are published on the websites of news channels, magazines, and agencies; these websites produce their news in more than one language, especially English and Arabic, and the translation is certified and official by these institutions and not a machine translation. The analytical methodology engages a comparative approach, adopting Toury's (1995) "problem-orientated approach", only sentences exhibiting translation gaps were analyzed. Accurate translations were excluded, as they do not contribute to identifying GPT's limitations in politically sensitive contexts. Such a selective method is in line with systematic functional linguistics (Halliday, 1985), where deviations in intellectual, interpersonal and textual descriptive functions are prioritized. The method begins with an English source text and its corresponding Arabic translation, as officially published by the relevant institutions. This sample is then processed using an AI translation tool, wherein the source text is inputted with a directive to generate an Arabic translation. The AI-produced output is subsequently analyzed in contrast to the institutionally authorized Arabic translation to identify variations in lexical, syntactic, and stylistic choices. This comparative framework allows for an assessment of the AI's alignment with or deviation from established translational norms.

The present study relies on GPT-40-mini—an AI language model developed by OpenAI. This version considers that at the time of writing, the study is considered one of the latest versions of GPT as an artificial intelligence tool, and there is no doubt that in each version the algorithms are improved to present results that are better than previous versions, so this version was chosen as an artificial intelligence model for the study.

**3.1 Data Collection:** The study relies on collecting data from official news websites (CNN, DW, BBC, Al Jazeera, and some other organizations) that produce their news in both English and Arabic, focusing on the Palestinian issue and the conflict in Gaza in particular as a case study. A huge amount

of data is translated, and translations that are too close to the human translations are filtered out; translations with gaps are taken, analyzed in tabular form, and compared with the approved human translation.

### Source text ST. 1:

“Three months after she came into this world Juri is set down on a slab in a Gaza hospital morgue drained of all signs of life the Palestinian Health Ministry says she is the 66th child to die of acute malnutrition during the war in Gaza. I do not know what the reason is and what their sin is to be killed in this manner. her father says their only fault is being innocent babies”. CNN (2025, June 27).

### Human translations (TT. 1):

"بعد ثلاثة أشهر من قدومها إلى هذا العالم، وضعت جوري على لوح في مشرحة أحد مستشفيات غزة، وقد استنزفت كل علامات الحياة. تقول وزارة الصحة الفلسطينية إنها الطفلة رقم ٦٦ التي تموت بسبب سوء التغذية الحاد خلال الحرب في غزة. لا أعرف ما هو السبب وما هي خطيئتهم ليقتلوا بهذه الطريقة! يقول والدها: "إن ذنبهم الوحيد هو أنهم أطفال أبرياء". (سي ان ان العربية، ٢٧ حزيران ٢٠٢٥).

### Chat GPT translation (TT. 2):

"بعد ثلاثة أشهر من قدومها إلى هذا العالم، وضعت جوري على لوحة في ثلاجة موتى مستشفى في غزة، وقد أزيلت جميع علامات الحياة. تقول وزارة الصحة الفلسطينية إنها الطفل السادس والستون الذي يموت من سوء التغذية الحاد خلال الحرب في غزة. يقول والدها: "لا أعرف ما هو السبب وما ذنبهم ليقتلوا بهذه الطريقة ذنبهم الوحيد هو أنهم أطفال أبرياء". (جي بي تي، ٣٠ حزيران ٢٠٢٥).

**Discussion:** In the analysis, human translation successfully integrates real reporting on institutional contexts—such as a Health Ministry report—with the profound emotional weight of the human tragedy, especially evident in the father’s hopeless appeal, including the poignant phrase "I do not know the reason". On one hand, the translation demonstrates high fidelity to both

the factual content and emotional tone of the source text (ST), preserving key details about Juri's death while effectively conveying the original's affective impact through culturally resonant Arabic phrases like "أطفال أبرياء" "innocent children". Conversely, GPT undermines the depiction of paternal despair by omitting the father's critical expression of uncertainty and utilizing overly literal phrasing that detracts from the emotional depth of the narrative. The source text (ST) strikes a careful balance between individual human tragedy and broader institutional reporting. HT successfully mirrors this dual focus by adhering to formal Arabic journalistic conventions while preserving an empathetic tone. It employs precise factual references, such as "الطفلة رقم ٦٦" (Child-66), alongside language that is suitable for a humanitarian narrative.

In contrast, TT2 disrupts this equilibrium by placing excessive emphasis on clinical details, using starkly literal terminology like "ثلاجة موتى" "dead body refrigerator". This choice introduces an inappropriate sense of detachment and conflicts with the text's necessary humanitarian sensitivity. Linguistically, the ST features lexically dense and syntactically complex constructions—such as "acute malnutrition" and "drained of life"—to convey the devastating effects of war. TT1 effectively naturalizes these elements into Arabic, employing equivalents like "سوء التغذية الحاد" "acute malnutrition" and "استنزفت كل علامات الحياة" "drained of all signs of life", while maintaining the ST's intricate syntax and causal linkages, as exemplified by "بعد ثلاثة أشهر" "after three months". This approach not only preserves the original meaning but also enhances the emotional resonance of the narrative in the target language. TT2, however, relies on overly literal and unnatural renderings, such as "أزيلت جميع علامات الحياة" "all signs of life were removed", which lack fluency and rhetorical impact, resulting in a divided structure.

In terms of register, the ST employs a formal yet emotive tone. TT1 achieves a culturally appropriate formal register that retains the necessary emotive

weight, using phrases like "ذنبهم الوحيد" "their only guilt/sin" to underscore the tragedy of innocent suffering. TT2, however, exhibits significant register inconsistency, incorporating clinical and insensitive terminology that diminishes the text's intended emotional and professional impact. Ultimately, TT1 exemplifies culturally adaptive translation (Toury, 2012, p. 78), preserving the ST's humanitarian ethos through natural Arabic syntax and an affective lexicon. TT2's strict literalism, on the other hand, undermines both functional and cultural frameworks, introducing inappropriate clinical terms while omitting key rhetorical elements, thereby weakening interpersonal engagement. This contrast underscores the limitations of artificial intelligence in translation, which, despite lexical accuracy, often struggles with pragmatic and cultural competence (Baker, 2018, p. 72). The following table demonstrates the analysis of the translations according to Halliday's systemic functional linguistics:

Table 1 Comparative Analysis of Metafunctions in Translations HT and GPT

Metafunction	TT1	TT2
<b>Ideational</b>	Clear transitivity (war → death)	Disjointed "أزيلت علامات الحياة"
<b>Interpersonal</b>	Empathetic (لا أعرف السبب)	Detached (omits father's voice)
<b>Textual</b>	Cohesive (...تقول الوزارة)	Incohesive "ووضعت على لوحة"

It seems clear that TT1 demonstrates a culturally adaptive approach, skillfully modifying both the emotional tone and factual content of the source text to align with Arabic linguistic and cultural norms. By employing culturally resonant phrases, it preserves the original's intent while ensuring natural readability for the target audience. In contrast, TT2 struggles with significant limitations, primarily due to its overreliance on literal translation. This approach not only diminishes its equivalence to the original text but also compromises readability. Additionally, TT2 violates standard Arabic

grammatical conventions—such as unnecessarily using passive constructions where active voice would be more appropriate (Baker, 2018, p. 67)—further weakening its coherence and stylistic effectiveness. The following table illustrates the analysis via Toury’s descriptive translation studies model.

Table 2 The Analysis via Toury’s Descriptive Translation Studies Model

DTS Element	TT1	TT2
<b>Initial Norm</b>	<b>Acceptability:</b> Natural Arabic	<b>Adequacy:</b> Overly literal
<b>Operational Norm</b>	Employs Arabic emotive conventions	Misuses clinical terms
<b>Acceptability</b>	High (matches TL norms)	Low (awkward phrasing)

In brief, the analysis highlights that TT1 outperforms TT2 in functional equivalence, readability, and lexical faithfulness. TT1 effectively preserves the emotional tone, cultural nuances, and syntactic flow of the source text, aligning with Arabic norms. In contrast, TT2 suffers from literal phrasing and fragmented syntax, reducing its acceptability and coherence.

### ST. 2:

Hajja Fatima, 73, sits silently at a mass grave containing her extended family: "They were playing hide-and-seek when the missile hit. I identify them by their shoes". **DW.** (2025, May 25).

### **TT.1:**

الحاجة فاطمة (٧٣ عاماً) تجلس صامتة عند مقبرة جماعية لعائلتها: "كانوا يلعبون الغميضة عندما أصابهم الصاروخ"، أعرفهم من أحذيتهم. (دي دبليو عربية، ٢٥ أيار ٢٠٢٥).

## TT. 2:

تجلس حجة فاطمة، البالغة من العمر ٧٣ عامًا، بصمت عند مقبرة جماعية تحتوي على عائلتها الممتدة: "كانوا يلعبون الغميضة عندما أصابتهم الصاروخ. أتعرف عليهم من أحذيتهم". (جي بي تي، ٢٧ حزيران ٢٠٢٥).

**Discussion:** Translating English texts into Arabic for informational purposes necessitates careful attention to linguistic accuracy, cultural sensitivity, and stylistic appropriateness. A comparative analysis of two target texts featuring Hajja Fatima's tragic account reveals significant variations in translation quality. TT1 effectively conveys the content with precision, accurately translating "Hajja Fatima" as (الحاجة فاطمة), thereby preserving the religious and cultural honorific implications associated with the term. This translation not only respects the cultural significance of "Hajja" but also maintains grammatical correctness. In contrast, TT2 incorrectly translates the term as (حجة فاطمة), introducing a spelling error and failing to capture the culturally appropriate form of address. This discrepancy highlights the importance of meticulous attention to detail in translation, particularly when dealing with culturally significant terms.

Furthermore, the focus on personal tragedy in the ST is retained in TT1, which effectively sustains the emotional weight of the narrative. TT2, however, weakens the cultural context by employing an incorrect lexical choice, thereby diminishing the text's authenticity. The genre of the ST, identified as a narrative report, is appropriately adapted in TT1 as an authentic Arabic narrative, whereas TT2 exhibits a mixed register, compromising stylistic coherence. Lexical precision is particularly critical in this context, as evidenced by the treatment of the religious and culturally specific term "Hajja". TT1 correctly employs (الحاجة), an honorific that aligns with Arabic conventions, while TT2 erroneously uses (حجة), a term that does not carry the same sociolinguistic significance. Additionally, TT2 contains a

grammatical error in feminine verb agreement (أصابتهم), further undermining its reliability. These discrepancies highlight the importance of meticulous translation practices to ensure fidelity to both linguistic structure and cultural nuance.

This aligns with Arabic cultural and linguistic norms (Toury, 2012, p. 63). In contrast, TT2 fails due to the use of (حجة), which means in Arabic, in the Islamic religion (Hajj/الحج), or a term for (evidence/الدليل) (not a religious or honorific term) and the feminine verb (أصابتهم), which incorrectly genders the masculine noun (الصاروخ). These errors violate basic Arabic grammar rules and cultural conventions, disrupting both functional equivalence and target-language norms. This preserves the ST's emotional gravity and fulfills Systemic Functional Linguistics' (SFL) metafunctions: ideationally by maintaining accurate processes, interpersonally through a respectful tone, and textually via cohesive narrative structure (Halliday & Matthiessen, 2014, p. 91).

The assessment of the translation using Halliday's systemic functional linguistics is illustrated in the following table:

Table 3 Grammatical Errors and Their Effects on Accuracy, Credibility, and Cohesion

Metafunction	TT1	TT2
<b>Ideational</b>	Accurate process representation	Distorted by grammatical error
<b>Interpersonal</b>	Empathetic, respectful	Loses credibility
<b>Textual</b>	Cohesive narrative	Disjointed due to errors

In examining the impact of translation errors, it becomes obvious that these mistakes significantly distort the interpersonal metafunction, compromising the authenticity of the text, and that reflects Toury's (2012, p. 79) concept of "negative interference," where "machine translation imposes inappropriate lexical and grammatical choices." The next table presents the analysis using Toury's approach of descriptive translation studies:

Table 4 Cultural Appropriateness Versus Grammatical Flaws in Arabic Localization

DTS Element	TT1	TT2
<b>Initial Norm</b>	<b>Acceptability:</b> Culturally accurate	<b>Adequacy:</b> Literal but flawed
<b>Operational Norm</b>	Uses culturally appropriate terms	Misapplies honorific and grammar
<b>Acceptability</b>	High (matches Arabic norms)	Low (cultural/grammatical errors)

Finally, the analysis underscores the necessity of employing culturally informed and grammatically precise translation strategies. TT1 exemplifies an effective approach by maintaining content accuracy, emotional resonance, and stylistic integrity, whereas TT2 demonstrates the pitfalls of inadequate lexical and grammatical choices. Such findings emphasize the need for rigorous quality control in translation to preserve the original text's meaning and impact.

### **ST. 3:**

"I held my dead baby for days': Gaza mother's trauma after airstrike." **BBC.** (2023, November 5).

## TT. 1:

احتفظت بجثة طفلي لأيام.. صدمة أم غزاوية بعد غارة جوية". (بي بي سي العربية، ٥ تشرين الثاني ٢٠٢٣

## TT. 2:

"حملت طفلي الميت لعدة أيام: صدمة أم من غزة بعد الضربة الجوية". (جي بي تي، ٢٤ حزيران ٢٠٢٥).

**Discussion:** This comparative analysis examines human (TT1) and machine-generated (TT2) translations of a journalistic text depicting a Gazan mother's psychological trauma after an airstrike. The study demonstrates how translation methodologies differentially affect semantic precision, cultural resonance, and stylistic coherence in the target text. TT1 exemplifies culturally adaptive translation (Toury, 2012), prioritizing fidelity to the source text's psychological and emotional dimensions. Lexical choices such as "احتفظت بجثة طفلي / kept the corpse of her child" preserve the narrative's focus on prolonged grief, while culturally anchored phrases like "أم غزاوية" / "a Gazan mother" enhance contextual relevance. The translation adheres to systemic functional linguistics (SFL) principles (Halliday & Matthiessen, 2014), maintaining ideational accuracy (e.g., nominalized "استنزفت كل علامات الحياة" - "drained all signs of life" and interpersonal engagement e.g., the rhetorical "ما هو السبب؟" - "what is the reason?").

Conversely, TT2 adopts a literal approach, introducing semantic and pragmatic distortions. The phrase "حملت طفلي الميت / carried my dead child" shifts emphasis from psychological retention to physical action, diluting the original's traumatic gravity. Register inconsistencies (e.g., mixing formal "أم

– "ثلاجة موتى" (with colloquialisms) and terminological missteps e.g., "morgue refrigerator" disrupt textual cohesion and journalistic tone. Such flaws reflect machine translation's tendency to prioritize formal equivalence over functional adequacy, violating descriptive translation norms (DTS) and SFL's metafunctional requirements. competence (Baker, 2018, p. 72). The next table illustrates the assessment of the translations based on Halliday's systemic functional linguistics:

Table 5 Comparative Analysis of Metafunctions in Translations HT and GPT

Metafunction	TT1	TT2
<b>Ideational</b>	Clear transitivity (war → death)	Disjointed "أزيلت علامات الحياة"
<b>Interpersonal</b>	Empathetic (لا أعرف السبب)	Detached (omits father's voice)
<b>Textual</b>	Cohesive (...تقول الوزارة)	Incohesive "وُضعت على لوحة"

This section explores the effectiveness of each translation in conveying emotional depth and factual accuracy within the cultural context. TT1 achieves a similar effect by changing the emotional and factual content of the original text to fit Arabic culture, using phrases that resonate culturally e.g., "أطفال أبرياء" and standard terms like "مشرحة". In contrast, TT2 does not succeed because it makes mistakes in literal translation, such as using "ثلاجة موتى" for "morgue," which affects both its similarity to the original and its readability. The machine translation also breaks Arabic grammar rules by using passive voice "وُضعت" when active voice would be better, making it less acceptable (Baker, 2018, p. 67). Conversely, TT2 falls short due to these literal translation errors, which disrupt both equivalence and readability. The next table presents the analysis using Toury's approach of descriptive translation studies:

Table 6 A Comparative Analysis of Norms in GPT and Human translation

DTS Element	TT1	TT2
<b>Initial Norm</b>	<b>Acceptability:</b> Natural Arabic	<b>Adequacy:</b> Overly literal
<b>Operational Norm</b>	Employs Arabic emotive conventions	Misuses clinical terms
<b>Acceptability</b>	High (matches TL norms)	Low (awkward phrasing)

Overall, TT1 is superior to TT2 in achieving functional equivalence and target-language appropriateness. Human translation effectively conveys the psychological trauma through culturally resonant Arabic journalistic conventions, while AI's literal approach misrepresents key semantic elements. TT1 adheres to Arabic media norms through proper lexical choices and consistent register, whereas TT2 suffers from terminological mismatches and stylistic inconsistencies.

#### **ST. 4:**

“Iran Warns of 'Regional War' If Gaza Offensive Continues” **Al Jazeera**,  
(2023 Oct. 15)

##### **1. TT. 1:**

إيران تحذر من "حرب إقليمية" إذا استمر الهجوم على غزة (الجزيرة، ١٥ تشرين الأول ٢٠٢٣)

##### **2. TT. 2:**

"ايران تحذر من "حرب إقليمية" إذا استمرت العمليات في غزة" (جي بي تي، ١٣ حزيران ٢٠٢٥)

The original text is a political statement conveying Iran's warning of a "regional war" if the military offensive in Gaza continues. Human translation precisely renders "offensive" as (الهجوم على غزة), maintaining the aggressive connotation and formal diplomatic register. In contrast, the machine translation (TT2) neutralises the original tone by using (العمليات في غزة), which misleadingly suggests a routine military operation rather than an aggressive attack. While TT1 aligns with Arabic media norms and preserves the accusatory tone, TT2 softens the rhetoric, distorting the intended meaning.

In connection with Halliday's metafunction; the following table demonstrates the analysis of the translations according to Halliday's systemic functional linguistics:

Metafunction	TT1	TT2
Ideational	Accurately represents violent action (يهاجم)	Neutralizes action (يعمل)
Interpersonal	Maintains authoritative urgency	Weakens credibility
Textual	Cohesive with political discourse	Disjointed from context

TT1 indicates full compliance with SFL standards by ideationally preserving the source text's aggressive semantic force through the accurate lexical choice (يهاجم), interpersonally maintaining the requisite diplomatic urgency, and textually aligning with Arabic public policy discourse. Conversely, TT2 exhibits systemic deficiencies; ideational meaning is compromised through the neutralised term (يعمل), interpersonal personal power is reduced by

stylistic weaknesses, and textual coherence suffers from register inconsistency. Human translators surpass accurately conveying agency, sustaining rhetorical impact, and ensuring contextual coherence—critical aspects that machine translation often struggles to replicate due to its inability to fully process pragmatic and ideological subtleties.

Toury's Descriptive Translation Studies (DTS) framework evaluates translations in relation to the norms of the target culture. On one side, the human translation (HT) demonstrates acceptability by aligning with the conventions of Arabic political discourse. On the other side, the GPT-generated translation tends toward adequacy but imposes a form of neutrality that appears unnatural within the target context. Human translation's operational norms are evident in its culturally appropriate lexical choices, which enhance its acceptability. Conversely, TT2's emphasis on neutrality leads to ideological distortion, ultimately diminishing its acceptability within the target culture, reducing its acceptability:

DTS Element	TT1	TT2
<b>Initial Norm</b>	Acceptability: Culturally accurate	Adequacy: Overly neutral
<b>Operational Norm</b>	Uses politically apt lexicon	Imposes unnatural neutrality
<b>Acceptability</b>	High (matches Arabic media norms)	Low (ideological distortion)

The analysis indicates full compliance with SFL standards by ideationally preserving the source text's aggressive semantic force through the accurate lexical choice (يهاجم), interpersonally maintaining the requisite diplomatic urgency, and textually aligning with Arabic public policy discourse. In contrast, GPT demonstrates several systemic shortcomings. The ideational meaning is weakened by the use of the neutralized term (يعمل), which fails to convey the intended nuance. Additionally, the interpersonal dimension of personal power is diminished due to stylistic inadequacies. Furthermore, the overall textual coherence is undermined by inconsistencies in register, resulting in a less effective translation. Human translators surpass accurately conveying agency, sustaining rhetorical impact, and ensuring contextual coherence—critical aspects that machine translation often struggles to replicate due to its inability to fully process pragmatic and ideological subtleties.

In summary, human translation accurately conveys the source text's aggressive tone and political urgency through precise lexical choices (الهجوم) that align with Arabic discourse norms, while machine translation neutralises meaning (العمليات) and weakens rhetorical impact. Halliday's SFL and Toury's DTS analyses reveal TT1's superiority in maintaining ideational, interpersonal, and textual functions, whereas TT2 distorts meaning through unnatural neutrality. That highlights machine translation's limitations in handling nuanced political discourse compared to human translators' contextual and cultural competence.

### 3.2 Findings

The study highlights distinct patterns in the types of errors observed. These errors fell into four main categories: lexical, syntactic, ideological, and cohesion-related issues.

Table 1 Comparative Analysis of Error Types Across Samples

Sample	Lexical Error	Syntactic Error	Ideological Bias	Cohesion Error	Total Errors per Sample
1	2	1	0	3	6
2	2	0	0	3	5
3	1	1	0	3	5
4	1	0	1	11	4
<b>Total</b>	<b>6</b>	<b>2</b>	<b>1</b>	<b>12</b>	<b>all samples=20</b>

These findings, analysed through the frameworks of Toury's Descriptive Translation Studies (DTS) and Halliday's Systemic Functional Linguistics (SFL), underscore that the most significant challenges in these translations pertain to cohesion and lexical accuracy, while syntactic and ideological issues appear to be less pronounced. The analysis reveals that errors related to cohesion are the most common of all identified errors. This finding highlights the significant challenges translators face in maintaining cohesive text structures, which are essential for effective communication and comprehension. This distribution indicates that challenges with textual cohesion are prominent in the studied corpus, although lexical and syntactic difficulties also play a significant role. The findings of this study reveal limitations in GPT-generated translations of sensitive geopolitical content from English to Arabic, particularly regarding cultural nuance and textual

coherence. This supports Hypothesis 1 regarding accuracy gaps but contradicts Hypothesis 2's assertion about the potential for error elimination. It also shows ideological bias. When analysed through Toury's Descriptive Translation Studies (DTS) framework, the lexical and syntactic errors were found to reflect violations of operational norms. Additionally, Halliday's Systemic Functional Linguistics (SFL) analysis highlighted systemic challenges related to cohesion within the textual metafunction. Collectively, these findings indicate that while GPT demonstrates basic grammatical fluency, it struggles with discourse-level coherence and cultural adaptation—critical elements for the accurate translation of politically sensitive content.

The study concludes that current AI translation systems require substantial improvements in managing contextual nuances and maintaining textual cohesion to achieve human-level quality in the translation of geopolitical news, particularly concerning contentious topics such as the Gaza conflict.

## References

1. Baker, M. (2018). *In other words: A coursebook on translation* (3rd ed.). Routledge.
2. Gadamer, H.-G. (1975). *Truth and method* (2nd ed.). Continuum. (Original work published 1960)
3. Garcia, M. (2021). Contextual ambiguity in human and machine translation. *Translation Studies Quarterly*, 34(4), 112–129. <https://doi.org/10.1080/13556509.2021.1956782>
4. Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.
5. Johnson, T., & Lee, S. (2022). Adaptability in translation technologies: A comparative analysis. *Journal of Language Technology*, 15(1), 45–63. <https://doi.org/10.1017/jlt.2022.8>

6. Kenny, D. (2021). Machine translation for everyone: Empowering users in the age of artificial intelligence. Language Science Press.
7. Koehn, P. (2009). Statistical machine translation. Cambridge University Press.
8. Manning, C. D., & Schütze, H. (1999). Foundations of statistical natural language processing. MIT Press.
9. Müller, K. (2023). Ethics in AI-powered translation: Bias and accountability. *Ethics & Technology*, 9(3), 201–220. <https://doi.org/10.1007/s10676-023-09703-z>
10. O'Hagan, M. (2016). The role of technology in translation. *The Translator*, 22(2), 127–145.
11. Toury, G. (2012). *Descriptive translation studies – and beyond* (Revised ed.). John Benjamins.
12. Williams, A. (2019). *The art of human translation: Tradition and practice*. Academic Press.