

How AI Translation Personas Reframe Diplomatic Texts

Mohd Nour Al Salem¹, Sharif Alghazo^{2,3}, Nimer Abusalim¹, Ghaleb Rabab'ah^{2,3}

¹Department of English Language and Literature, School of Foreign Languages, The University of Jordan, Jordan

²Department of Foreign Languages, Faculty of Arts, Humanities and Social Sciences, University of Sharjah, UAE

³Department of English Language and Literature, The University of Jordan, Jordan

Correspondence: Sharif Alghazo, Department of Foreign Languages, College of Arts, Humanities and Social Sciences, University of Sharjah, UAE; Department of English Language and Literature, The University of Jordan, Jordan.

Received: November 24, 2025

Accepted: January 6, 2026

Online Published: January 19, 2026

doi:10.11114/smc.v14i2.8391

URL: <https://doi.org/10.11114/smc.v14i2.8391>

Abstract

This study explores how persona prompting in AI translation systems, particularly OpenAI's ChatGPT (GPT-4o), shapes politically sensitive discourse. By assigning translation personas such as "diplomat" and "politician," the AI produced distinct Arabic renderings of six controversial statements by Prince Zeid Ra'ad Al Hussein, former UN High Commissioner for Human Rights. Drawing on Descriptive Translation Studies, Critical Discourse Analysis, and Narrative Theory, the analysis examines tone, word choice, and rhetorical strategies. Findings show that AI translations not only transfer meaning but also reframe ideas: diplomats soften charged language, while politicians balance criticism and formality. These shifts highlight translation as a political and ethical act with implications for global perceptions, policy, and public opinion. The opacity of AI processes further raises concerns over transparency, accountability, and bias. The study contributes to research at the intersection of computational linguistics, translation studies, and political discourse, urging critical scrutiny of how AI translation shapes meaning and ideology in international communication.

Keywords: AI translation, persona prompting, diplomatic discourse, narrative framing, ChatGPT, translation ethics, rhetorical analysis

1. Introduction

Rapid advancements in artificial intelligence (AI) translation technology have transformed cross-linguistic communication. Recent developments in large language models (LLMs), such as GPT-4 (OpenAI, 2023), enable AI systems to translate text and modify its style, tone, and intent through persona prompting. Persona prompting involves giving the AI specific instructions to translate from the perspective of a particular role, such as a diplomat or neutral observer. This creative method adds new layers to translation, especially in sensitive areas, such as diplomacy, where the wording and tone of messages can carry significant political weight.

Diplomatic language refers to "the official oral remarks or written documents, including external speeches, statements, communiqués, treaties, notes and letters, etc., given by the state leaders or officials representative of the government stand, on diplomatic occasions" (Yang, 2012, p. 2). Translation in diplomatic language may not be neutral; it influences negotiations, international relations, and conflict resolution through word choice. Barston (1997) argues that diplomatic communication contributes "to the pacific settlement of disputes between involving states and/or other actors" (p. 215). Therefore, ambiguous language is often deliberately used in diplomatic statements to balance competing interests or convey subtle messages to different audiences. When AI translation systems apply personas that highlight specific viewpoints or emotional tones, the way political comments are interpreted globally can change. For example, a "nationalist diplomat" persona may emphasize issues of sovereignty and security, while a translation using a "peace advocate" persona might soften contentious phrases. These differences indicate that AI-driven persona prompting can actively reinterpret diplomatic discussions rather than simply reproducing information from the source language.

As globalisation and digital diplomacy grow, understanding the impact of AI translation personas is increasingly important. Governments, media outlets, and international organizations increasingly rely on AI translation systems to communicate quickly and across multiple languages. The political and rhetorical implications of persona prompting remain poorly understood in machine translation research, which primarily focuses on improving language accuracy and fluency and reducing bias (Belinkov & Glass, 2019). Human translators make interpretive choices shaped by cultural,

ideological, and ethical factors, as translation studies have pointed out (Venuti, 2012).

This study is justified both theoretically and practically. Effective communication is crucial for international diplomacy, making it essential to understand the role of AI personas to avoid unforeseen diplomatic consequences. The study extends the concepts of fidelity and equivalence to encompass AI's evolving role, thereby advancing translation studies. Nida's foundational theory distinguishes between dynamic equivalence, which focuses on the message's impact on the audience, and formal equivalence, which aims for literal accuracy (Nida, 1964). An aspect of this dynamic is persona prompting, in which AI systems using preprogrammed personas actively select among conflicting equivalence strategies. This approach reframes translation as an algorithmic act of role-playing that mediates and influences political discourse. To examine the specific linguistic and discursive changes brought about by persona prompting and their broader diplomatic implications, the main question for this study would be:

How do AI translation personas influence political and diplomatic statements concerning rhetorical strategy, word framing, and emotional tone?

2. Literature Review

Previous translation studies have examined the translation of religious texts from Arabic into English (e.g., Al Salem et al. 2025; Al Salem et al. 2023; Burohaima et al, 2025a; Burohaima et al, 2025b) and other Arabic linguistic structures (e.g., Al Salem et al. 2024). Machine Translation technology evolved from relying on rules and statistics to the neural machine translation (NMT) approach, the predominant method employed nowadays. Initial rule-based systems, such as SYSTRAN, relied on rigid language rules and matching expressions, resulting in translations that tended to sound stilted and ignore context. NMT improved translation by leveraging deep learning and vast amounts of bilingual text. Deep Learning and Google Translate are two examples that employ an encoder-decoder architecture to produce improved context-aware translations. Research has shown that the fluent, smooth quality of neural machine translation (NMT) is sometimes deceptive. Ustaszewski (2019) concluded that NMT tends to produce well-fluent text, but fluency sometimes suppresses errors of adequacy, such as meaning insertions or omissions. Errors of the latter kind unless the translation is closely compared with the original text, may be missed by readers, as by readers, and post-editors.

At the same time, advancements in large language models have introduced new ways to build output with persona prompting. Asking a model to take on some role, such as diplomat, reporter, or neutral observer, allows end-users to guide the response's tone and perspective. Hu and Collier (2024) found that role instructions of this sort actually influence how models construct and interpret information, but only to a relatively small extent. Their study demonstrates how small changes in the persona can alter the nuance of information conveyed by machines, especially in delicate or interpretive scenarios. Framing theory is a useful framework for examining how meaning changes during transitions, particularly when strong political emotions are involved. Later developed by Entman (1993) as the option and salience of scenes of reality, framing theory helps clarify how language shapes perception.

Studies in translation studies (e.g., Baker, 2006; Cheikh and Rabab'ah, 2025) have shown that translators necessarily frame through lexical choices, syntactic reorganization, and cultural mediation. When AI translation systems are promoted to "perform" in specific roles, they contribute to the same framing procedures, often without clear awareness or accountability. For instance, instructing a model to translate as a "peace negotiation" can yield more pacific language than a prompt to translate "from a military perspective", even when both are interpretations of the exact source text.

Empirical research also finds that the same biases present in society may be reflected in and even magnified in the outputs of AI systems. Caliskan et al. (2017) created the Word Embedding Association Test (WEAT) to show that many standard word embeddings, including those trained using GloVe or word2vec, reproduce implicit biases from people. For instance, the models would associate female words more with family and male words with career, more so than was typically associated in society at large.

Equally, Prates et al. (2020) proved that Google Translate often lacked male pronouns when translating from gender-neutral languages. These findings question AI translation accuracy. In high-stakes environments like diplomatic negotiations, the risks are combined. Translations can subtly shift the intent or perceived aggressiveness of a statement, particularly if the AI system has been prompted with a specific role. Bender et al. (2021) believe that large language models are "stochastic parrots", repeating models from data without real understanding, which makes them especially exposed to imitating dominant ideological narratives. In persona-driven translation, the "role" played by the AI can act as a filter, expanding or minimizing specific aspects of the original message.

The intersection of translation studies, political discourse, and artificial intelligence (AI) is a complex and urgent area of research. Given advancements in machine translation and its widespread use in diplomacy and politically sensitive contexts, the framings of translated messages and, particularly, how prompts, which assign roles or personae to the AI system, impact this, also become increasingly relevant. In doing so, this review situates the study within three significant literature domains: translation theory relevant to political texts, nascent work on AI's production of political language, and the paucity of literature on persona-based prompting in AI translation practices.

The translation of political and diplomatic texts contributes to the field of Translation Studies at the juncture of ethical responsibility, ideological framing, and socio-political stakes (see Alshakhanbeh & Alghazo, 2022; Benyakoub et al., 2022). Functionalist approaches to translation place strong emphasis on the fact that the translator's strategies ought to be guided by the purpose—or Skopos—in the text. The concept was initially outlined by Hans Vermeer (1978), who asserted that translation is always purposeful and that strategies must be correlated to communicate. Katharina Reiss (1989) followed this with text typologies based on overriding functions (e.g., informative, expressive, or operative), which became the cornerstone for how the translator would proceed. Role-based instructions in modern-day use with AI translation can be considered the digital equivalent of thinking in terms of Skopos. Just as the human translator would adjust their strategy depending on whether the text is intended to persuade, inform, or inspire, the AI program can be instructed to "translate this like a diplomat" or "like a journalist." These instructions nudge the model toward a desired endpoint in communication, with role-based prompting as the latter-day equivalent of Skopos-driven strategy in human translation. Here, the AI is encouraged to follow specific communicative intentions, such as neutrality or advocacy, simulating human translator decisions.

Toury's (1995) descriptive translation studies model posits the concept of translation norms, predisposed rules of translation that are enforced in specific cultures or institutions, in his overall model of how translations are produced, circulated, and consumed. Toury Differentiates between initial norms (which determine whether to stay closer to the source or target culture) and operational norms (which guide decision-making during translation). These norms have immediate consequences for politically motivated translation: diplomatic translations may tend towards domestication to avoid offending, whereas activist translations may opt for foreignization to draw attention to an ideological position. Persona prompts do so implicitly in AI translation, where each persona corresponds to idiomatic, ideologically, and institutionally mediated stances.

The translator's ethical positioning (Baker, 2006) also applies to AI-brokered translation. Baker says that translators cannot help but have their part to play in framing the narratives of conflict, often entrenching or reworking perceptions. Because AI models imitate personas, they also participate in narrative creation. When the translator's agency is influenced by personal, political, or institutional connections, the AI's agency is constrained by prompt design and training data.

The increasing pervasion of AI in language services has led to a greater questioning of the impact of ideological content on these systems. Recent research suggests that large language models (LLMs) systematically replicate the information on which they are trained and produce the biases and political narratives embedded in the training data. One such point is AI's limited ability to find idiomatic and nuanced political expressions. Alkhawaja (2024), however, concluded that LLMs struggle with idiomaticity and fine-grained pragmatic meaning, even when they produce grammatically sound output. These difficulties may diminish the rhetorical power of diplomatic cables or activist statements, which frequently use metaphor, euphemism, or culturally specific language.

Obeidat and Jaradat (2024) focus on translating Arabic resistance literature, considering AI's effectiveness. Their findings help illustrate both the strengths and weaknesses of systems like ChatGPT. Though AI sometimes benefits from its ability to generate syntactically harmonious translations, it frequently fails to preserve elements such as tone, cultural referents, or political insinuations found in the original texts. This is especially the case in resistance literature, which typically uses emotive and ideologically 'hot' language. These problems are compounded when using translations as input to unsupervised general-purpose models, without a persona-specific prompt.

Law similarly illuminates how AI processes esoteric political speech. Regarding legislative texts, ElFqih and Monti (2024) study ChatGPT's overall quality when translating Arabic legislative texts, finding difficulties in keeping terminological accuracy and contextual consistency. This is even more reason to justify control as to the forcing of the issue, particularly when translating texts that politically matter. Alkhawaja (2024) compares ChatGPT-3 with human translators and shows that ChatGPT-3 can perform well on simple tasks, though it benefits from prompt engineering. This illustrates how persona prompting helps move from generic production to rhetorically sensitive translation.

Taken together, these works demonstrate that AI is making steps as a tool for translation within a political lens, but its effectiveness in a political context is strongly reliant on context-aware prompt generation, subtle modeling of ideological framing, and post-editing by human editorial users. However, despite this progress, the framework for initiating AI translations has not received extensive investigation. While much of the previous work naturally emphasizes technical quality, fluency, accuracy, and coherence, little prior work has addressed how a persona assigned to an AI translator (such as a diplomat or a politician) affects the strategic framing of the text.

Persona prompting is an interesting intersection of NLP and translation studies. Instead of surface-level style adjustment, it shapes the manner in which the text is aligned rhetorically, politically, and relationally with its reader. Most recent work by He (2024) is telling: if you prompt ChatGPT specifically as a translator, the result is far more effective than prompting as an "author" or with no persona at all. Such evidence would suggest that role-based prompting can profoundly influence the style and communicative intent of a translation.

Likewise, experiments with Role-Playing Language Agents (RPLAs) illustrate how activating personas in AI systems (professional roles or population profiles) shapes the tone, framing, and orientation of their outputs (Chen et al., 2024). By the same token, persona prompting may be regarded as more than a technical device: it is a strategic device for the simulation of communicative intents with specific application in sensitive areas like diplomacy. However, the in-the-world effects of such prompts on readers, as well as their utility for diplomatic ends, remain underdeveloped topics for future research.

Present prompt-engineering experiments hardly affect the stylome and rhetome, which have different effects on the reader's attention. Yamada (2023) highlights the potential for rapid engineering in low-resource language pairs, though it does not examine any political or ethical issues. In politically fraught environments, these subtleties are essential. The persona of the translator (or the AI) can mediate if the text soothes conflict or sharpens it. It is time to explore how personality-based AI translation services work as an analogy toward or in interference with conventional translation regulations, especially in high-stakes diplomatic situations. This study seeks to fill that gap by systematically demonstrating how AI adaptive translations change in tone, style, and framing depending on context-specific prompts, all of which inform both translation theory and the broader ethical dialogue on AI in international communication.

3. Methodology

This study employs a qualitative comparative research design based upon Descriptive Translation Studies (Toury, 1995), Critical Discourse Analysis (Fairclough, 1995), and narrative theory for studies in translation (Baker, 2006). The project explores the effect of role-prompting of an AI language model (e.g., ChatGPT) upon the translation of politically charged texts. It examines the influence of different instructed personas—diplomat or politician—upon the reinterpretation of the linguistic, ideological, and affective framing of translations. The case study is conducted using six of Prince Zeid Ra'ad Al Hussein's high-profile public addresses, a Jordanian diplomat and former United Nations High Commissioner for Human Rights, known for using morally and diplomatically charged language.

Data collection

Six inflammatory public statements by Prince Zeid were selected for translation into the corpus. They were selected for their politically divisive nature, rhetorical impact, and thematic relevance to diplomacy, conflict, and human rights. Examples are the description of Syria as a "torture chamber," the condemnation of Security Council vetoes, and the call for the cessation of Israel's occupation of Palestine. The words were collected from publicly available, reliable sources, including NPR, The Guardian, and UN documents. The six source texts were translated from English into Arabic using the OpenAI ChatGPT-4 model GPT-4o under two prompting conditions:

1. "Translate this accurately into Arabic as if it were produced by a diplomat."
2. "Translate this accurately into Arabic as if it were produced by a politician."

These prompts were designed to replicate different rhetorical functions, potentially with different ideological and emotional framings. All translations were stored and formatted the same across all the outputs.

Organizing and Comparing Outputs

The machine translations were placed into a parallel corpus. The different versions of the same statement were presented in tables to permit comparison. Alignment made it possible to monitor the shifting of tone, word selection, and ideological position in role-based instructions of different types. By conducting a DTS-based comparative analysis following Toury's Descriptive Translation Studies model, the translations were examined for overt regularities and divergences in tone, lexical selection, and narrative stance. Lexical signals such as the employment of affectively unmarked words (e.g., "conflict" as opposed to marked ones like "slaughterhouse"), register variation, and the use of mitigating or accusatory words were recorded in particular. The focus was on observing patterns in relation to the impact of the AI persona framing on the process and norms of translation.

Critical Discourse Analysis (CDA)

To investigate the ideological dimensions of the translations, a critical discourse approach (Fairclough, 1995) was employed. All translations were analyzed in order to determine:

- The register and tone used (strategic, formal, and emotional)
- The portrayal of agents and victims
- The narrative perspective (accusatory, advocacy-based, neutral)
- The ideological stance inferred from any role-based directive

The interpretation relied upon the assumption that diplomats like balance and legalism in their language, whereas politicians combine pragmatism with strategic vagueness. Comparing the patterns with the AI translations allowed us to assess to what extent the model reproduces or exaggerates these types of discourse. Each AI-translated response was

then compared with what actual diplomats and politicians tend to say in the same given context. The objective was to examine whether the tone, framing, and ideological stance of the AI-generated response for each role-play prompt match or align with actual samples of the varied roles. This provided a grounded analysis of the way AI role-play prompts represent and distort actual discursive practices. All the texts used for the research were in the public domain. No confidential information or personal data was processed. Prince Zeid's identity was used purely rhetorically and academically for stylistic simulation analysis and not impersonation. The AI model was executed under human supervision, and all interpretations were cross-verified to ensure no misrepresentation.

4. Analysis and Findings

As previously stated, this paper aims to explore how AI algorithms (namely, ChatGPT using its GPT-4 model) produce translations of statements deeply rooted in politically contentious contexts. All six examples employed in this research are English statements produced by Prince Zeid on several controversial political topics. Each statement is translated by ChatGPT into Arabic using two different prompts:

1. "Translate this accurately into Arabic as if it were produced by a diplomat."
2. "Translate this accurately into Arabic as if it were produced by a politician."

Upon analyzing the results, it is evident that ChatGPT's role-playing algorithms mold each statement to conform to the norms, tone, and emotiveness of the designated Target Language "speaker" in both scenarios, and therefore succeed in assuming each persona.

Below is a detailed analysis of the translations of each statement:

1. Syria as a "Torture Chamber"

Original Statement (English): "Syria is now a torture chamber. A place of savage horror and absolute injustice. Surely there must be accountability for such monstrous crimes."

Persona	Arabic Translation	Back Translation
Diplomat	"تعد سوريا حالة فريدة من حيث قسوة ووحشية هذا الصراع، والمعاناة التي لحقت بملايين المدنيين الذين لا يستحقون هذا المصير. لقد أصبحت البلاد بأكملها كأنها غرفة تعذيب."	Syria is a unique case in terms of the cruelty and brutality of this conflict and the suffering that affected millions of civilians who do not deserve this fate. The whole country is now as if it were a torture chamber.
Politician	"إن ما يحدث في سوريا يتميز بوحشية لا مثيل لها، وقد تسبب هذا النزاع في معاناة لا تطاق لملايين المدنيين. أصبحت البلاد بمثابة 'غرفة تعذيب شاملة'."	What is happening in Syria is marked by unmatched brutality. This conflict has resulted in unbearable suffering for millions of civilians. The country has become a full-scale torture chamber.

The Arabic translations of Prince Zeid's metaphor calling Syria a "torture chamber" show the different ways in which the message changes based on the translator's role. The diplomatic version uses polite and formal language, describing Syria as a "unique case" of cruelty and suffering. The use of the simile "كأنها غرفة تعذيب" ("as if it were a torture chamber") rather than the definitive declaration in the original softens the English metaphor, making it sound more like an observation than a direct accusation. This fits diplomatic norms, which favor careful, respectful language to avoid intense conflict, following Toury's Descriptive Translation Studies idea of a "norm-governed shift" in which translations are adapted to fit cultural expectations. The political translation is more direct, using decisive expressions like "وحشية لا مثيل لها" ("unmatched brutality"), "معاناة لا تطاق" ("unbearable suffering"), and describing Syria as "غرفة تعذيب شاملة" ("a full-scale torture chamber"). This clearly highlights the severity of the crisis while still maintaining a formal, composed tone. Drawing on Fairclough's (1995) Critical Discourse Analysis, it is clear that this version criticizes the Syrian regime indirectly yet firmly, balancing condemnation with political caution.

2. Security Council Vetoes and "Slaughterhouses"

Original Statement (English): "Some of the permanent members of the Security Council have repeatedly used the veto to block action... leaving the slaughterhouses of Syria and other conflict zones untouched."

Persona	Arabic Translation	Back Translation
Diplomat	"تستخدم صلاحيات الفيتو في مجلس الأمن أحياناً لحماية المصالح الوطنية، حتى في ظل وقوع فظائع جماعية، مما ساهم في خلق بعض من أسوأ بؤر العنف في عصرنا."	The veto powers in the Security Council are sometimes used to protect national interests, even when mass atrocities are taking place. This has contributed to the creation of some of the worst hotbeds of violence in our time.
Politician	"لقد استخدم الفيتو في مجلس الأمن لحماية المصالح السياسية للدول الكبرى رغم وقوع فظائع مروعة، ما أدى إلى نشوء مجازر بشرية غير مسبوقة في السنوات الأخيرة."	The veto in the Security Council has been used to protect the political interests of major powers despite horrific atrocities taking place. This has led to the emergence of unprecedented human massacres in the last years.

The translations above reflect different tones shaped by the roles they play. The diplomatic version uses soft language and avoids direct blame. It uses a present simple passive structure to refer to a general truth: ("تستخدم صلاحيات الفيتو")

("Veto powers are used"), instead of the more direct original, which, albeit unspecifically, still points the finger at some permanent members of the Security Council for using the veto on specific occasions, additionally, by using "ساهم في خلق" ("contributed to the creation of some of the worst hotbeds of violence"), the translation tones down the harsh "slaughterhouses" metaphor. The phrasing also suggests that the veto power is "sometimes" ("أحياناً") used with the justified aim of "حماية المصالح الوطنية" (protecting national interests), whereas the original clearly states otherwise. This fits the diplomatic style, which tries to express concern without causing confrontation. This follows Toury's idea that cultural and professional norms shape translations, and, in this case, diplomacy's preference for a cautious, neutral tone. The political version maintains a close imagery, but adds more explanation. It states that the veto powers had been used to "protect the political interests of major powers" (حماية المصالح السياسية للدول الكبرى), which "has led" to "مجازر بشرية غير مسبوقة" ("unprecedented human massacres"). This version still clearly criticizes the system, but it presents the issue in a more calculated, policy-driven manner typical of political speech.

3. Israel's Occupation of Palestine

Original Statement (English): "Ending the occupation is essential for peace and would benefit both Palestinians and Israelis."

Persona	Arabic Translation	Back Translation
Diplomat	"الإبقاء على الاحتلال لن يؤدي إلا إلى استمرار المعاناة الشديدة للطرفين. وإنهاءه قد يفتح الباب أمام الفلسطينيين والإسرائيليين لإيجاد السلام."	An ongoing occupation will only lead to continued immense suffering for both sides. Ending it might open the door to peace for both Palestinians and Israelis.
Politician	"استمرار الاحتلال يعني مزيداً من الألم للطرفين، بينما يمثل إنهاءه فرصة حقيقية لتحقيق السلام بين الشعبين الفلسطيني والإسرائيلي."	An ongoing occupation means more pain for both sides, while ending it represents a real opportunity to achieve peace between the Palestinian and Israeli peoples.

The Arabic translations of Prince Zeid's statement highlight how tone and framing shift based on the speaker's persona. The diplomatic version uses cautious language, such as "الإبقاء على الاحتلال لن يؤدي إلا إلى استمرار المعاناة" (An ongoing occupation will only lead to continued immense suffering), which softens the call to action. It uses cautious language, suggesting that if the occupation were to end, peace would be only a possibility, not an inevitable outcome: "وإنهاءه قد يفتح الباب للسلام" (ending the occupation might open the door to peace). This follows Toury's idea that translators adjust their choices to fit the norms of diplomatic communication, which is calm, respectful, and open to dialogue. The political version is more direct and definitive, stating that "ending the occupation represents a real opportunity to achieve peace": ("يمثل إنهاءه فرصة حقيقية لتحقيق السلام"). It balances moral pressure and political realism, a common theme in political discourse. From a Critical Discourse Analysis perspective, this version manages power and responsibility, suggests change, and leaves room for negotiation.

4. Trump Administration's Immigration Policy

Original Statement (English): "The thought that any state would seek to deter parents by inflicting such abuse on children is unconscionable."

Persona	Arabic Translation	Back Translation
Diplomat	"فكرة أن تلجأ أي دولة إلى ردع الآباء من خلال إلحاق هذا النوع من الأذى بالأطفال هي أمر يثير القلق العميق."	The idea that any state would resort to deterring parents by inflicting this kind of harm on children is a deeply troubling matter.
Politician	"من غير المقبول أن تستخدم أي دولة إيذاء الأطفال كوسيلة لردع الأسر عن الهجرة، فهذا السلوك يتنافى مع القيم الإنسانية الأساسية."	It is unacceptable for any state to use the abuse of children as a means to deter families from migrating, as this behavior contradicts fundamental human values.

The translations of Prince Zeid's criticism of the Trump administration show how language can shift depending on the speaker's role. The diplomatic version says it is "deeply troubling", "أمر يثير القلق العميق", to harm children to deter parents, rather than calling it "unconscionable." This more careful tone reflects Toury's (1995) view that translators work within cultural and institutional norms. Diplomats often avoid strong judgment and prefer gentler phrasing to keep discussions open. The political version keeps the criticism clear, but it also adds a layer of explanation. It unapologetically states that it is "unacceptable" and "contradicts fundamental human values", rather than just "deeply troubling", but also talks about "abuse of children as a means to deter families", "إيذاء الأطفال كوسيلة لردع الأسر", which frames the issue more as a policy failure than as moral outrage. This reflects Fairclough's (1995) idea that political discourse manages blame strategically. It criticizes without sounding too emotional, and allows room for negotiation or reform.

5. Myanmar's Treatment of Rohingya as a "Textbook Example of Ethnic Cleansing"

Original Statement (English): "Because Myanmar has refused access to human rights investigators, the current situation cannot yet be fully assessed, but the situation seems a textbook example of ethnic cleansing."

Persona	Arabic Translation	Back Translation
Diplomat	نظرًا لرفض ميانمار السماح بدخول محققين في مجال حقوق الإنسان، لا يمكن تقييم الوضع الحالي بشكل كامل، ولكن يبدو أنه يمثل حالة نموذجية للتطهير العرقي.	Given Myanmar's refusal to allow human rights investigators entry, the current situation cannot be fully assessed, but it appears to be a textbook case of ethnic cleansing.
Politician	رفض ميانمار دخول محققين لحقوق الإنسان يعيق التقييم الشامل للوضع، إلا أن المؤشرات تشير إلى أن ما يحدث يُعد مثالًا واضحًا على التطهير العرقي.	Myanmar's refusal to allow human rights investigators entry hinders a comprehensive assessment of the situation; however, indicators suggest that what is happening constitutes a clear example of ethnic cleansing.

In his original statement, Prince Zeid was careful with his words. He stated that the situation “seems a textbook example” of ethnic cleansing and noted that investigators did not have complete access. The diplomat’s Arabic version keeps this cautious tone by using “يبدو أنه يمثل” and “لا يمكن تقييم الوضع بشكل كامل”. This soft, careful language follows diplomatic norms, in which it is important not to accuse a country without confirmed evidence. According to Toury’s (1995) approach, this reflects how translations often follow the expectations of the setting, and, in this case, those of a UN official who needs to be balanced and professional. The politician’s version uses firmer language. It states that it is indeed a “clear example of ethnic cleansing”, “مثالًا واضحًا” instead of employing the more cautious “seems to be,” referring matter-of-factly to official “indicators”: “تشير المؤشرات إلى أنه مثال واضح على التطهير العرقي” (“Indicators suggest that it is a clear example of ethnic cleansing”). It also plainly points out Myanmar’s refusal as the main impediment for assessment: “رفض ميانمار دخول محققين لحقوق الإنسان يعيق التقييم” (“Myanmar's refusal hinders assessment”). This version still sounds official but clearly blames Myanmar to a greater degree than its diplomatic counterpart. From Fairclough’s (1995) view, the political voice here is more direct and holds responsibility, while still staying within formal language. It fits political speech where leaders are expected to speak out, but carefully.

6. Criticism of Philippine President Duterte’s War on Drugs

Original Statement (English): “It makes one believe that the president of the Philippines needs to submit himself to some sort of psychiatric evaluation.”

Persona	Arabic Translation	Back Translation
Diplomat	قد يدفع هذا البعض إلى الاعتقاد بأن رئيس الفلبين قد يستفيد من إجراء تقييم نفسي معين.	This might lead some to believe that the Philippine President could benefit from undergoing a certain psychiatric evaluation.
Politician	تصرفات رئيس الفلبين قد تثير تساؤلات بشأن حالته النفسية، وربما تستدعي إجراء تقييم طبي مناسب.	The Philippine President's actions might raise questions about his mental state, perhaps warranting a proper medical assessment.

Zeid's original statement is a clear criticism of President Duterte's approach to the war on drugs, suggesting the need for psychiatric evaluation. By employing “might lead some to believe that” (“قد يدفع البعض إلى الاعتقاد بأن”), the translation implicitly and subtly negates the idea. Also, the phrasing “could benefit from a psychiatric assessment” (“قد يستفيد من”), instead of the original “needs to” demonstrates how the diplomat's translation softens this critique by utilizing a softer tone which is more indirect and polite. This fits diplomatic language norms that avoid direct criticism and allow room for interpretation. Following Toury’s DTS, this version respects diplomatic norms by softening the claim to avoid confrontation and maintain formal relations. The politician’s translation is firmer, using “تصرفات رئيس الفلبين قد تثير تساؤلات بشأن حالته النفسية” (the actions of the Philippine President might raise questions about his mental state), and suggesting that a medical evaluation may be necessary. This reflects political speech that is more direct yet calculated. It openly questions the president’s mental state. However, it remains within the formal and respectful bounds, and it is consistent with Fairclough’s idea that political discourse balances criticism with strategy to maintain authority.

As can be observed from the above analysis, each AI translation shifts slightly in impact, tone, and word choice for each statement, which shows how translations can reflect and shape the message depending on the purpose, and the speaker’s social and political role - careful in diplomacy and strategic in politics. In other words, the diplomat softens and the politician balances. These shifts show that translation is not just about words, but also about who is speaking, why, and how their story is being told.

6. Discussion

The present study contributes to the growing body of research at the intersection of translation studies, political discourse, and artificial intelligence by foregrounding the interpretive role of persona-based prompting in AI-mediated translation. Rather than treating machine translation as a neutral or purely technical process, the discussion advances the argument that persona prompting activates discursive, ideological, and ethical dimensions that closely resemble—yet also complicate—those traditionally associated with human translators. Drawing on insights from Descriptive Translation Studies (Toury, 1995), Critical Discourse Analysis (Fairclough, 1995), narrative theory (Baker, 2006), and recent scholarship on large language models (Bender et al., 2021; Hu & Collier, 2024; Chen et al., 2024), this section interprets the broader implications of the study’s outcomes without reference to individual textual cases.

One of the central theoretical implications of this study is the reaffirmation of translation as an act of framing rather than mere linguistic transfer. Framing theory, as articulated by Entman (1993), emphasizes that communicators select certain

aspects of reality and make them more salient in order to promote specific interpretations, evaluations, and responses. Translation—particularly political and diplomatic translation—has long been recognized as a site where framing is enacted through lexical choice, syntactic restructuring, modality, and evaluative language (Baker, 2006; Venuti, 2012).

The introduction of persona prompting into AI translation systems intensifies this framing function. By instructing an AI model to translate “as a diplomat,” “as a politician,” or any other role, the user effectively pre-selects a communicative stance, thereby shaping how agency, responsibility, and affect are distributed in the target text. This process aligns closely with Baker’s (2006) argument that translators are narrators who reconfigure events and actors through narrative positioning. What distinguishes AI-driven persona prompting is that this narrative positioning is not the result of lived experience, institutional affiliation, or ethical deliberation on the part of the translator, but rather of algorithmic patterning guided by prompt design and training data.

From a theoretical standpoint, this finding supports the view that AI translation systems participate in framing not despite their lack of consciousness, but precisely because they operationalize framing as a probabilistic selection process. As Bender et al. (2021) caution, large language models reproduce dominant patterns embedded in their training data, effectively amplifying conventional discourses under the guise of neutrality. Persona prompting, therefore, does not introduce subjectivity into an otherwise objective system; instead, it makes explicit the ideological filtering inherent in AI-generated language.

Within the framework of Descriptive Translation Studies, Toury’s (1995) concept of norms provides a productive lens for interpreting the effects of persona-based translation. Toury distinguishes between initial norms (orientation toward source or target culture) and operational norms (decisions made during the translation process). In traditional human translation, these norms are shaped by professional training, institutional expectations, and sociopolitical context. In AI translation, persona prompts function as surrogate norms, guiding the system toward particular regularities in tone, register, and evaluative stance.

Importantly, these norms are not embedded permanently within the system but are activated dynamically through user input. This flexibility challenges conventional understandings of normativity in translation. Whereas Toury conceptualizes norms as relatively stable and socially regulated, persona prompting introduces a fluid, user-driven normativity that can shift from one translation to the next. This raises critical questions about accountability: if translation norms are instantiated through prompts, responsibility for ideological framing moves upstream—from the translator to the prompt designer.

This shift has significant implications for politically sensitive translation. Diplomatic and political texts are traditionally governed by highly institutionalized norms that prioritize caution, ambiguity, and face-saving strategies (Barston, 1997; Yang, 2012). Persona prompting simulates these norms effectively, but without the ethical reflexivity that human translators bring to high-stakes contexts. As a result, AI translations may appear institutionally appropriate while simultaneously obscuring the interpretive decisions that shape them.

Nida’s (1964) distinction between formal and dynamic equivalence remains influential in discussions of translation purpose and audience effect. Dynamic equivalence prioritizes the target audience’s response, often at the expense of formal correspondence with the source text. Persona prompting in AI translation can be understood as a computational instantiation of dynamic equivalence: the system is guided to produce an output that aligns with the perceived communicative goals of a particular role.

However, unlike human translators, AI systems lack an understanding of audience reception beyond statistical approximation. While they can generate linguistically and pragmatically plausible texts, they cannot evaluate the ethical consequences of prioritizing certain audience effects over others. This limitation is particularly salient in political discourse, where shifts in tone or emphasis can influence public opinion, policy debates, and international relations.

The study thus invites a reconsideration of dynamic equivalence in AI contexts. If equivalence is achieved through role-based simulation rather than interpretive judgment, then equivalence itself becomes a function of prompt engineering. This raises the possibility that AI translations may optimize for rhetorical appropriateness while undermining transparency and fidelity at a deeper ideological level. As Venuti (2012) warns, fluency and naturalness can mask the translator’s intervention; in AI translation, this masking effect may be even more pronounced. A recurring concern in AI research is the reproduction of social and political biases through algorithmic systems (Caliskan et al., 2017; Prates et al., 2020). The present study extends this concern to the domain of persona-based translation, where ideological bias is not only inherited from training data but also actively shaped by user-defined roles. This dual source of bias complicates simplistic narratives about AI neutrality.

Critical discourse analysis emphasizes that power operates through discourse by normalizing certain perspectives while marginalizing others (Fairclough, 1995). Persona prompting can reinforce this process by privileging institutionalized voices—such as diplomacy or governance—over alternative narratives. While such privileging may be appropriate in certain contexts, it also risks reproducing hegemonic discourses at the expense of dissenting or marginalized viewpoints. Moreover, because persona prompting is often framed as a stylistic feature, its ideological consequences may go unnoticed by end users. This aligns with Baker’s (2006) observation that narrative framing is most effective when it appears natural and unmarked. In AI translation, the smoothness and coherence of the output can obscure the fact that alternative framings were possible—and perhaps ethically preferable.

Recent work on role-playing language agents (Chen et al., 2024) conceptualizes personas as mechanisms through which AI systems adopt consistent behavioral and discursive patterns. When applied to translation, this role-playing capacity positions AI as a narrative agent that does more than transmit meaning: it actively constructs stories about events, actors, and responsibilities. This narrative agency challenges traditional distinctions between author, translator, and mediator. In persona-based AI translation, the system occupies an ambiguous position that combines elements of all three roles. It does not originate content, yet it reshapes it in ways that influence interpretation. It lacks intentions, yet it produces intention-like effects through prompt-conditioned output.

From a narrative theory perspective, this reinforces Baker's (2006) claim that narratives are not tied to individual consciousness but circulate through discursive practices. AI systems become new sites through which narratives are reproduced, stabilized, or subtly transformed. The ethical question, therefore, is not whether AI "intends" to frame discourse, but how its framing effects can be anticipated, evaluated, and governed.

The findings of this study have important implications for translation ethics. Traditional ethical frameworks in translation emphasize accuracy, impartiality, and accountability (Venuti, 2012). Persona prompting complicates these principles by introducing a layer of intentionality that is distributed across human users and machine systems. If a translation is shaped by a persona prompt, who is ethically responsible for its framing—the AI, the user, or the institution deploying it? Baker (2006) argues that translators are ethically accountable for the narratives they help circulate. Extending this argument to AI translation suggests that users and institutions must assume responsibility for prompt design and deployment. This is particularly critical in diplomatic and political contexts, where translations can have real-world consequences. From a professional standpoint, the study underscores the need for AI literacy among translators and language professionals. Rather than viewing AI as a replacement for human expertise, it should be approached as a powerful but potentially opaque tool that requires critical oversight. Training programs should equip translators to interrogate AI outputs, understand the effects of persona prompting, and make informed decisions about when and how to use such tools.

The ability of ChatGPT to mimic different voices in translation has important implications. First, there is an ethical responsibility that translators must recognize. AI tools do not simply translate words; they also reshape the message. For example, choosing a prompt such as "translate as an activist" may not be neutral, but a political choice that can change how the original speaker's intent is understood (Baker, 2006). Second, audience reception varies depending on the tone. A diplomatic style may be appropriate in formal contexts, while an activist tone could be more effective in engaging the public. Third, translators, including AI systems guided by prompts, are never neutral. The person giving instructions to the AI bears the responsibility for how the message is framed.

The results show that, when given proper instructions, ChatGPT can reproduce different rhetorical styles, whether the careful language of diplomacy or the give-and-take of political negotiation. This means that AI tools can do more than translate words; they can actually challenge our thinking and question the idea that translation is neutral. At the same time, these powerful features require us, as human users, to be careful and critical and to take ethical responsibility in using AI in translation. For translation pedagogy, these findings underline the need to train future translators and interpreters not only in linguistic competence but also in AI literacy. Students should critically evaluate AI outputs, understand how persona prompting influences framing, and develop strategies to ethically integrate machine translation into their practice. Classroom activities could include comparative analysis of persona-driven outputs, encouraging students to reflect on how role-based prompting alters meaning, tone, and political stance.

7. Conclusion

This study examined AI translation personas, specifically those created with advanced models such as ChatGPT, to reinterpret diplomatic statements in different contexts. It found that AI translations act as ideological agents that can alter messages by shifting perspective and framing. The study also explored how tone, word choice, and rhetorical techniques in these translations facilitate cross-language communication. These changes could drastically affect how people globally understand them, shaping public opinion, diplomatic responses, and policy discussions. One other important finding in this study is that, particularly when dealing with politically sensitive material, AI translations often use a softer, more impartial tone. This raises significant concerns about the validity and integrity of the message, even though it might help reduce conflict. The study also found significant distinctions between the political and diplomatic tones that go beyond stylistic variance and reveal more profound ramifications for how translated language conveys accountability and authority. In simple terms, the diplomat adopts a more measured tone, while the politician maintains a balanced stance. The finding that translation results differ depending on the prompt framing is interesting, as it highlights the importance of the interface and context in shaping AI responses based on the details provided.

More research is needed to examine how ongoing reliance on AI translation might affect long-term trends in diplomatic communication. Research could also investigate how AI translation styles may differ based on language pairs, such as English-Arabic and English-French. Further study may be needed to understand how different prompt structures affect AI

translation results. This includes various aspects of analysis to determine how explicit a persona or political views influence the final translation's neutrality or bias. This study underscores the need to understand AI's abilities in high-stakes communication settings to explore how human translators interact with AI outputs and whether these partnerships make diplomatic translation easier. As translation technologies continue to evolve, the intersection of artificial intelligence, linguistic agency, and global discourse will always require a thorough, multidisciplinary examination. Beyond its implications for diplomacy, this study also offers practical guidance for translator training. By incorporating persona prompting into translation pedagogy, educators can better prepare students to engage critically with AI systems, equipping them to identify bias, maintain ethical responsibility, and use AI strategically in professional practice.

Authors contributions

Dr. MNA and Dr. SA were responsible for study design and revising. Prof. NA and GR was responsible for data collection, and data analysis. MNA wrote the first draft and all authors revised it.

Funding

This work did not receive any fund.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Redfame Publishing.

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

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

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

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

References

- Al Salem, M. N., Alghazo, S. M., Alrashdan, I., Abusalim, N., & Rayyan, M. (2023). On English translation variation of similar plural nouns in the Holy Quran. *Cogent Arts & Humanities*, 10(1), Article 2196136. <https://doi.org/10.1080/23311983.2023.2196136>
- Al Salem, M. N., Aljadid, R., Alghazo, S. M. & Rayyan, M. (2024). Translation of cognate accusative from Arabic to English: The case of *Call of the Curlew*. *Cogent Arts & Humanities*, 11(1), Article 2377422. <https://doi.org/10.1080/23311983.2024.2377422>
- Al Salem, M. N., Altakhaineh, A., Alghazo, S. M., & Rayyan, M. (2025). On the translation of attributes of Allah in the Holy Quran: A comparative study. *Forum for Linguistic Studies*, 7(1), 232-243. <https://doi.org/10.30564/fls.v7i1.7393>
- Alkhawaja, L. (2024). Unveiling the new frontier: ChatGPT-3 powered translation for Arabic–English chat language pairs. *Theory and Practice in Language Studies*, 14(2), 347-357. <https://doi.org/10.17507/tpsl.1402.05>
- Alshakhanbeh, S., & Alghazo, S. M. (2022). A pragmatic analysis of criticism strategies against government policies on social media in Jordan: A gender-based investigation. *Jordan Journal of Modern Languages and Literatures*, 14(2), 263-285. <https://doi.org/10.47012/jjml.14.2.3>
- Baker, M. (2006). *Translation and conflict: A narrative account*. Routledge. <https://doi.org/10.4324/9780203099919>
- Barston, R. P. (1997). *Modern diplomacy* (2nd ed). Essex: Longman.
- Belinkov, Y., & Glass, J. (2019). Analysis methods in neural language processing: A survey. *Transactions of the Association for Computational Linguistics*, 7, 49-72. https://doi.org/10.1162/tacl_a_00254

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623). Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445922>
- Benyakoub, N., Alghazo, S. M., Altakhaineh, A., & Rabab'ah, G. (2022). A cross-cultural analysis of disagreement strategies in Algerian and Jordanian Arabic. *Kervan: International Journal of African and Asiatic Studies*, 26(1), 253-275.
- Burohaima, L.; Rabab'ah, G.; Benmessaoud, S. (2025a). Rewriting Women in the Qur'an: Gender, Ideology and Translation. *Open Cultural Studies*, 9(1), 20250080. <https://doi.org/10.1515/culture-2025-0080>.
- Burohaima, L.; Rabab'ah, G.; Benmessaoud, S. (2025b). Translating Arabic allusions into English: A hermeneutic approach. *Babel*, 71(2), 170–193. <https://doi.org/10.1075/babel.24091.bur>.
- Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183-186. <https://doi.org/10.1126/science.aal4230>
- Cheikh, M., & Rabab'ah, G. (2025). From the American sitcom “Seinfeld” to Modern Standard Arabic: Implicature translation by Arabic EFL students. *Ampersand*, 14, 100213. <https://doi.org/10.1016/j.amper.2024.100213>.
- Chen, J., Wang, X., Xu, R., Yuan, S., Zhang, Y., Shi, W., ... & Xiao, Y. (2024). From persona to personalization: A survey on role-playing language agents. arXiv preprint arXiv:2404.18231. <https://arxiv.org/abs/2404.18231>
- ElFqih, K., & Monti, J. (2024). Large language models as legal translators of Arabic legislation: Do ChatGPT and Gemini care for context and terminology? In *Proceedings of the Second Arabic Natural Language Processing Conference* (pp. 111-122). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.arabicnlp-1.10>
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51-58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. Longman.
- He, S. (2024). Prompting ChatGPT for translation: A comparative analysis of translation brief and persona prompts. arXiv preprint arXiv:2403.00127.
- Hu, T., & Collier, N. (2024). Quantifying the persona effect in LLM simulations. *arXiv preprint arXiv:2402.10811*. <https://doi.org/10.18653/v1/2024.acl-long.554>
- McGreal, C. (2017, June 6). UN rights chief calls for end to Israeli occupation of Palestinian territories. *The Guardian*. <https://www.theguardian.com/world/2017/jun/06/un-human-rights-council-israel-occupation-palestine>
- Mirzayev, T. (2025). AI translation in diplomacy. *The Review of Diplomatic Communication*.
- Nida, E. A. (1964). *Toward a science of translating: With special reference to principles and procedures involved in Bible translating*. Brill. <https://doi.org/10.1163/9789004495746>
- Obeidat, M. M., & Jaradat, M. A. (2024). Artificial intelligence accuracy in translating resistance literature from Arabic into English: Google Translate and ChatGPT as a model. *Research Journal in Advanced Humanities*, 5(2), 45-62.
- OpenAI. (2023). ChatGPT (GPT-4) [Large language model]. <https://openai.com/chatgpt>
- OpenAI. (2023). GPT-4 technical report. <https://cdn.openai.com/papers/gpt-4.pdf>
- Prates, M. O. R., Avelar, P. H. C., & Lamb, L. C. (2020). Assessing gender bias in machine translation: A case study with Google Translate. *Neural Computing and Applications*, 32(10), 6363-6381. <https://doi.org/10.1007/s00521-019-04144-6>
- Reiss, K. (1989). Text types, translation types and translation assessment. In A. Chesterman (Ed.), *Readings in translation theory* (pp. 105-115). Finn Lectura.
- The Guardian. (2017, June 6). UN human rights chief calls for end to Israel's occupation of Palestine. *The Guardian*. <https://www.theguardian.com/world/2017/jun/06/un-human-rights-council-israel-occupation-palestine>
- Toury, G. (1995). *Descriptive translation studies and beyond*. John Benjamins. <https://doi.org/10.1075/btl.4>
- Ustaszewski, M. (2019). Exploring adequacy errors in neural machine translation with the help of cross-language aligned word embeddings. In *Proceedings of the Human-Informed Translation and Interpreting Technology Workshop (HiT-IT 2019)* (pp. 107-114). Varna, Bulgaria. https://doi.org/10.26615/issn.2683-0078.2019_015
- Venuti, L. (2012). *The translator's invisibility: A history of translation* (2nd ed.). Routledge.
- Vermeer, H. J. (1978). Ein Rahmen für eine allgemeine Translationstheorie. *Lebende Sprachen*, 23(3), 99-102. <https://doi.org/10.1515/les.1978.23.3.99>
- Yamada, M. (2023). Optimizing machine translation through prompt engineering. *arXiv preprint arXiv:2308.01391*. <https://doi.org/10.48550/arXiv.2308.01391>
- Yang, M. (2012). The principles and tactics on diplomatic translation: A Chinese perspective. *Babel*, 58(1), 1-18. <https://doi.org/10.1075/babel.58.1.01min>