



AI Translation of Spoken Arabic: Human Bias and Quality Assessment

Ruba Alyaseen ¹

Research Scholar

Department of Linguistics

University of Kerala

Kerala

India

Mail ID: alyaseen.ruba@gmail.com

Mobile: 9778749550

Dr. S A Shanavas ²

Professor (retd)

Department of Linguistics

University of Kerala

Kerala

India

Mail ID: sasnavs@gmail.com

Mobile: 9495074811

APA Citation:

Alyaseen, R., & Shanavas, S. A. (2025). AI Translation of Spoken Arabic: Human Bias and Quality Assessment, *Journal of Indian Languages and Indian literature in English*, 02(04), 147-166; 2025

Submission Date: 12.03.2025

Acceptance Date: 24.03.2025

Abstract

Studies on AI bias and the ways to mitigate it are plenty, but what about human bias in the face of AI translation? As AI translation tools get more sophisticated and their role in mediating cross-linguistic communication increases, especially in multilingual contexts, understanding how native speakers perceive and judge AI outputs becomes essential. This study aims to investigate how accurately AI translates spoken Arabic (with its dialectical variations) and how human bias might shape such evaluation. This research uses a mixed-method approach to combine quantitative accuracy assessment and qualitative insights from native speakers of Arabic. It also investigates whether preconceptions about AI might impact translation assessment. This research

sheds light on the relationship between technological advancement and human perception. The results of this study contribute to the ongoing discussions concerning the evolving role of AI in translation, revealing further understanding regarding the reliability of machine translation as perceived by humans.

Keywords: Translation, AI translation, Human perception, Bias, Arabic, Spoken Arabic

1. Introduction

AI-powered tools nowadays have developed in such a way that they regularly generate human-like outputs, and people have started incorporating AI in all facets of life. Translation was incorporated into AI research and development as a main cross-linguistic and cross-cultural means of communication and an essential proponent of today's globalised world, due to the growing demand for quick and accurate translations. Translation between any language pair naturally endures some challenges, especially if the pair in question are from distinct language families with varying syntactic divergence and morphological complexity, as is the case with Arabic and English. Translating between these two languages is already an arduous task for human translators, so it is no wonder that Machine Translation (MT) and AI translation would also face some challenges in rendering accurate translations.

There is extensive research on translation between Arabic and English, but the Arabic in research so far has been mostly Modern Standard Arabic (MSA). Arabic is curious in the fact that its formal, written form and its spoken form are significantly different. As a well-known example of diglossia, its formal variety (taught in schools and used in written formal communication) differs linguistically from its spoken variety (acquired natively and used in all verbal communication) (Zbib et al., 2012). Such spoken variety of Arabic is referred to as dialect and the Arab world has that in abundance, due to varying geographical and socio-economic conditions of native speakers. While MSA is still present in all formal types of communication, it is dialectal Arabic that governs

informal day-to-day interactions, both spoken and written. Now media in all forms uses spoken Arabic (SA) to communicate with all types of audiences, and in order to reach further audiences globally, translation is a necessity. The literature is, however, lacking when it comes to translating SA. The norms regulating translation between MSA and English do not apply to translating SA. With the increasing use of AI, and most interactions being in dialectal varieties, research into AI reliability in translating SA is much needed.

And while it is true that AI output is at times treated with contempt, it is still part of today's consumed content. Taking into consideration that AI uses feedback from users to modify its output, a look into the perception of humans providing this feedback is worthwhile. AI offers significant benefits in cross-linguistic communication; that is why it is crucial to address the ethical concerns accompanying its use, like bias in this case. Insights on this aspect of AI research would lead to fair, responsible, and reliable translation practices.

1.1. Literature review

The sophistication of AI models and using advanced language models have helped AI-powered tools produce content that closely resembles human writing, making its detection increasingly difficult (Nguyen et al., 2023; Fraser, 2024). Humans' inability to detect AI-generated translations is a complex and multifaceted issue that is often hindered by flawed heuristics in the face of sophisticated language models (Jakesch, 2023). This struggle in discerning AI-generated translations can foster scepticism towards MT technologies, influencing users' perception in a way that they might question the authenticity and reliability of the content (Sadasivan et al., 2023).

The complex issue of audience bias towards AI translation encompasses various aspects, such as algorithm bias, cultural and sociopolitical implications, and ethical concerns. As AI becomes further integrated into communication, displaying such biases could impact audience perception of AI content. Much of the existing literature on AI translation examines AI bias. For quality assessment, it is a valuable area of study indeed, as some AI tools have been noted to display gender and racial stereotypes when translating, due to biased data they were trained on. Prates et al., (2020)

mention how Google Translate defaults to male pronouns in STEM-related job descriptions in a gendered context. And as Escudé Font (2019) iterates, such biases can instil harmful stereotypes by misrepresenting gender roles or inaccurately reflecting demographic distributions.

A study by Shahmerdanova (2025) highlights some AI translation challenges including cultural sensitivity and ethical concerns in various contexts and their role in affecting audience perception and acceptance of AI-generated translation. Caliskan (2023) stresses how implicit biases (regarding gender, race, ethnicity, etc.) can influence human assessment of AI-generated translations, and suggests that statistical regularities in language corpora might be to blame for affecting perceptions and interpretations of AI-generated content. According to Ji et al. (2023), Neural Machine Translation (NMT) is liable for misleading users by exhibiting cultural and linguistic biases and altering the interpreted meanings to fit the conventions and expectations of the target language audience, regardless of accuracy.

In a study by Zhu et al. (2024), 'human ego' was identified as a cognitive bias impacting assessment, where individuals prefer human-generated content due to their unwillingness to acknowledge AI's superiority in some creative and intellectual endeavours. This bias greatly affects perceptions of AI performance in creative writing contexts.

Zaki and Ahmed (2024) posit that biases present in training data can lead to biased translations in AI systems. As such, addressing audience bias demands research focused on identifying and mitigating these biases to guarantee culturally sensitive and inclusive translations across various contexts. Li (2025) points out how AI often struggles with translating cultural and linguistic nuances, leading to mistranslations or hollow translations. This in turn can affect human detection and perception due to the misrepresentation of cultural connotations, which ultimately diminishes the authenticity and effectiveness of crosslinguistic communication. This supports what Wells (2022) earlier expressed about how AI mistranslations of cultural and linguistic nuances can obscure context, leading to misinterpretations and loss of meaning, thus affecting how individuals detect and perceive the intent behind translated content.

Such issues will only be aggravated when there is no human oversight in AI translation since AI tools might not account for sociopolitical dynamics that are inherent in language use (Wells, 2022). Addressing audience bias towards AI translation, Deokar et al. (2024) suggest integrating Explainable AI (XAI) techniques as an attempt to mitigate biases, enhance transparency, and foster trust in AI-powered tools and AI-generated translations.

Quality-wise, studies demonstrate that in some contexts, AI-generated translations can match or exceed human translations, but they still fall short when dealing with humour and cultural references, thus impacting audience enjoyment and comprehension (Abu-Rayyash, 2024). The reception of AI tools and AI-generated content differs across linguistic communities, with varying perceptions influencing how such translations are trusted (Murayama et al., 2024). Studies reveal how human raters usually exhibit a strong preference for texts labelled as ‘human-generated’ over those labelled ‘AI-generated’, with a preference score exceeding 30% (Zhu et al., 2024), indicating a deep-seated scepticism toward AI content.

1.2. Research importance

The literature is rife with research on both translation quality assessment and AI translation bias, but the way this study approaches these two topics, entwining them and inferring new insights, makes for a novel way of considering the matters at hand. With the enunciated differences between MSA and SA, and the fact that the rules and the norms for translating one are naturally not the same for translating the other, this surely calls for comprehensive research regarding the different aspects of translating SA. This is where the importance of this research stems from. In attempting to test the quality of the output of the current prominent AI-powered translation tools, this study builds on existing research and furnishes further insights regarding the translation of dialectical varieties of Arabic using AI.

On the other hand, while the matter of AI translation bias has been discussed in abundance, human bias in the face of AI translations has not been studied enough. Bias against AI-created content already exists, so would bias against AI-generated translation be any different? This is what

the study aims to find out. People have started using AI to translate all types of texts and for all types of purposes. This reliance, while not always well rewarded, is sometimes also not well-perceived. The evaluation of AI-generated translations can be significantly influenced by various cognitive biases that affect human judgment, and such biases can lead to irrational preferences or perceptions regarding the quality of AI-generated translations compared to human-translated ones. In order to mitigate this way of AI-translation perception, and in order to examine whether it is even deserved, this study seeks to uncover any existing human biases toward AI translation. Understanding these biases, if they do exist, is crucial for improving human-AI collaboration and the overall acceptance of AI technologies.

1.3. Research questions

This study aims to ascertain two main points, as mentioned in the title: bias and quality in AI translation of SA. Regarding quality, this research attempts to answer whether AI translation of SA is up to par, which AI-powered translation tool handles this the best, and whether dialect plays a role in how well this is achieved. As for bias, in this case, the research means to find out whether or not preexisting notions about AI affect the assessment of translation (i.e. human good/AI bad, or vice versa). The answers to these questions will undoubtedly propel research on AI and translation forward.

1.4. Limitations

Due to constraints of time and resources, the quality assessment and the following survey were each limited to the same ten excerpts. Any more than this amount of texts would have put survey completion at risk, as it would have been too lengthy, tiring, and time-consuming for participants. Text-wise, the source language of the texts was Arabic, and while dialect choice was as varied as possible, this still was not enough to amount to all the spoken dialects of Arabic. And so, text choice was limited to the Levantine, Gulf, and Egyptian dialects. The translation was into English only and was made using the following AI-powered translation tools: ChatGPT, DeepL, Google Translate, and DeepSeek. The assessment part was handled by the researchers and no other experts

were consulted.

1.5. Methodology

This paper uses a mixed-method approach to get both qualitative and quantitative data that can further enrich the field with the resulting insights. Seeing as the aim of this paper is twofold, i.e. to assess which of the prominent AI-powered translation tools available these days is most helpful in translating SA in its dialectal varieties, and to ascertain whether human bias affects such quality assessment; the methodology used in this study is sectioned into two main parts as well. One part is related to quality assessment and one part is concerned with the human perception in relation to this assessment.

First, to test the quality of AI-generated translations, several variations were considered regarding text choice (i.e. language, dialect, number, length, and topic) and AI-powered translation tools. Ten short excerpts of SA were chosen from various media sources (including sitcoms, travelogues, and daily vlogs, to name a few) whose creators belong to various Arab nationalities and speak different dialects of Arabic. The reasoning behind this choice is that such media content -being advertised globally on social media- belies the need for further global outreach and some level of mutual understanding in the case of Arabic speakers. Thus, the excerpts chosen were ones that are usually understood by any Arabic speaker, making it, in turn, easier to distinguish any mishaps if mistranslated. The variety of dialects should help in assessing whether AI-powered translation tools can distinguish and handle each dialect. The texts were chosen from various topics in order to verify whether the context affects the quality of translation. While a more comprehensive list of texts would have been preferable and richer in data, due to the use of these texts in the second part of this study in a survey, it was best to contain the study within the suggested excerpts.

Each text was translated using the following AI-powered translation tools: ChatGPT, DeepL, Google Translate, and DeepSeek, since they are currently the most sophisticated and continuously updated. Based on established research, these tools were expected to handle Arabic translation well, even if in this case it is SA translation. And in case they did not fare well in translation, then the

ensuing results would still be used as a stepping stone into researching exactly in which areas AI translation is lacking.

Using intuitive assessment and comparative judgment, the best translation of each set was set aside to be used in the survey. As SA does not necessarily follow the rules of MSA, the way the translations were assessed was more along the lines of readability, naturalness, accuracy, and word choice. The other translations were rated best to worst afterwards, giving a proposed hierarchy of best AI-tools in translating SA, thus concluding the quality assessment part of the test.

As for the second part of the test, which deals with the possibility of human bias in regards to AI translations, the method used was a survey that was shared with native speakers of Arabic. The participants were chosen using the nonprobability snowball sampling method, starting locally with the Arabs studying at the University of Kerala, who in turn shared the survey form with their acquaintances as they saw fit, to reach Arabs studying in other Indian universities. As the texts used are of a variety of Arabic dialects, and the participants also belong to a variety of Arab countries, a section at the start of the survey enquires about the nationality and the dialects recognized by each participant, to see if such background has any effect on bias or assessment.

In order to check whether there were preexisting biases towards AI, a section of the survey clearly asks the participants whether they have used AI to translate before, whether they can spot AI translations, and what they think of such translations. And in order to check for any changes in perception post-test, a section at the end of the survey checks again whether their views on AI translations have changed after reviewing the translations.

In the assessment section of the survey, the participants were asked four questions in relation to each text: 1) to rate the accuracy of the translation, 2) to rate the naturalness of the translation, 3) to choose the main issues they found in the translation (if any exists), and 4) to guess whether the translation was made by AI or a human translator. What the participants did not know, of course, was that all the translations were AI-generated. In fact, they were the aforementioned ‘best translations’ picked through the assessment in the first part of the test. The fact that some of the AI-

generated translations were very human-like helped the matter. However, while hiding such crucial information was necessary in order not to affect the way participants answer, thus guaranteeing the most authentic data possible, it was still not ethical. To solve this, the introduction of the survey only vaguely indicated the purpose of the study being related to AI translations as perceived by humans, and a section was added at the end of the survey (post submission) indicating what the purpose of the survey really was. A note was left for the participants to email the researchers in case they feel curious about the answer results.

With the texts chosen, translated, and rated; and with the survey formed and administered, the data was collected and analysed to be discussed in the next section.

2. Discussions

While assessing the quality of the AI-generated translation, a noticeable level difference was observed immediately. DeepL and Google Translate usually fare well when it comes to Arabic-to-English translation, but in the case of SA, there were many cases of mistranslation. In the case of Text 8 (سمبوسة ومسخن في اكلة واحدة يبقى اكيد اللي اخترعها فلسطيني لما فكر خارج الصحن), DeepL even disregarded half of the text. It gave the translation “*It must have been invented by a Palestinian who didn't think outside the box*” when the translation should have been more along the lines of [Sambousa and Musakhan in one dish? It must have been a Palestinian who came up with this by thinking outside the ‘plate’!].

For the most part, there was an apparent shortening of sentences and vague translation that somewhat blurred the original meaning. For example, in Text 9 (قلي حاضر بس ما طلعتش حاضر لأي حاجة), which can be literally translated as [He said okay/ready, but he wasn't ready for anything], involving a play on words, DeepL translated it as “*Say yes, but never say yes to anything*” and Google Translate translated it as “*He said yes but he didn't say yes to anything*”. Both translations are way off the mark here that only the pronoun and the negation survived the translation. The resulting translations lack any sort of meaning or relation to the source text.

Similarly, in Text 2 (عاملي فيها أبو العريف), which translates to [He's acting like a know-it-all], both

DeepL and Google Translate gave such outlandish translations (“*A worker for me in Abu Al-Arif*” and “*He treated me like a chief*”, respectively) that the resulting output hardly seems like a translation any more.

On the other hand, both ChatGPT and DeepSeek were neck to neck in accurately translating the texts, having almost identical translations in many cases. ChatGPT only had a slight edge on DeepSeek in keeping the register of the texts in translation more apparent. Word choice-wise, DeepSeek definitely ventured out to change some words to suit a logical output in English. Whereas ChatGPT kept close to the original, even attempting a casual form of translation to preserve the feel of the source text. For instance, in Text 7 (طلعوا اذكيا ستاتنا لما شنشلو حالن بالذهب بالزمنات) which translates to [*Our grandmothers were quite smart when they draped themselves with gold back in the day*], DeepSeek chose the term “*invested in gold*” while ChatGPT went with “*decked themselves out in gold*”, the latter being closer to the correct meaning and register of the source text.

That does not mean that all the translations were accurate, however. All AI-powered translation tools seemed to fumble with Text 6 (يعني اليوم شو مشكل صار بالشغل وقلن انا ماخصني. انا بشو وانتي بشو), which can be translated as [*I mean what a problem happened at work today! And I kept telling them it had nothing to do with me. Where am I and where are you?*]. While all the translations were grammatically correct and gave the general meaning of “*a problem happened at work that wasn’t my concern*”, the last part of the text was mistranslated in all instances and some inferences about it were made in translation to suit the output regardless of what it meant in the source text. The last part of the sentence should have been translated along the lines of “*see the difference between what I’m talking about and what you’re on about*” or the literal translation “*where am I and where are you (in regards to the topic)*”. AI-generated translations, however, were “*What does that have to do with me and you?*”, “*I am what I am and you are what you are*”, “*What’s up with me and what’s up with you?*”, and “*What’s it to me, and what’s it to you?*”. Admittedly, this part can be particularly tricky even when translated by a human, so the way it was mistranslated by AI was probably a given.

As for the survey, and across 50 answers, the results were mostly consistent. First, in the pretest section, most of the participants answered that AI-generated translations are somewhat accurate. They had high confidence in accurately spotting AI-generated translations where all answers were either “very confident” or “moderately confident”. But at the end of the survey, this confidence seems to shake a little. When asked post-test if their confidence in identifying AI-generated translation has changed, 48% of the answers were “maybe” and 28% were “yes”. The remaining 24% that answered “no” seems a bit ironic when recalling that all the text translations were AI-generated. Some of the participants added notes in the “other” section to mention that they double-checked AI to see if it *‘really could translate dialects this well’*.

Not knowing in hindsight that all the translations were AI-generated, participants picked “human” as the translator over 55% of the time in three cases (namely texts 1, 3, and 7). Meanwhile, in the cases of texts (2, 8, 9, and 10), the results were $\approx 50\%$ for both AI and human translations. The majority of participants selected AI translation for texts (4, 5, and 6). Out of the original ten texts, which were all AI-translated, only these three texts were agreed upon by the majority of over 60% of participants as AI-translated.

In survey questions 2 and 3 (following each text translation), the participants were asked to rate the naturalness of the translation and then to pick what the main issues were with the translation if any existed. When the participants picked “human” for the translator, they always picked “very natural” or “somewhat natural” and “no issues”. But incidentally, even when they selected “AI”, they still mostly went with “somewhat natural” and “no issues”, i.e. responses that can be considered a nonbiased assessment. In cases where the participants found fault in the translations, they cited “lack of cultural/contextual nuances”, “misinterpretation” and “unnatural language/word choice” regardless of whether they have selected AI or human as the translator. This can also be taken as a sign of unbiased assessment. Even the previously mentioned 24% of participants, who maintained they could distinguish AI-generated translation from a human-translated one when they clearly could not, have still fairly evaluated the translations.

3. Conclusions

Using the results discussed earlier, several inferences can be deduced. While AI translation of SA is not up to par when compared to that of MSA, AI-powered translation tools are still developed enough to distinguish the various dialects and come up with appropriate, mostly accurate translations. They were not all at the same level, of course. ChatGPT delivered the most acceptable and human-like translations, and DeepSeek was a close runner-up. Google Translate barely managed to outrank DeepL as the latter gave the most disappointing and inaccurate responses. Further research can delve into the reasons behind such abysmal results compared with the outstanding output of ChatGPT and DeepSeek.

Dialect-wise, the Levantine dialect (especially if the context is humour) proved most difficult to translate, as indicated by AI tools misinterpreting the meanings and thus by survey participants pointing out the ensuing issues. Further research into the translation of the Levantine dialect (into English or otherwise) can shed more light on what the cause of such problematic translations can be. Furthermore, other dialects of spoken Arabic should be put to the test in the case of AI translation to enhance and better generalise translation universals.

Concerning bias, while participants were not as good as they believed at detecting AI content, they were still, for the most part, unbiased in translation assessment. The fact that they branded some translations as AI-generated and some as human-translated did not affect further evaluation. If the translation sounded natural and had no issues, it was reported as such. Likewise, when participants disapproved of a translation, they mentioned any and all issues they recognised. Background experience with AI and native nationality or dialect did not affect the evaluations, and neither did the aforementioned ‘human ego’. Further research can look for the reasons that lead to labelling one translation as AI-generated and one as human-translated.

This study was an attempt to enrich our understanding of the ever-evolving field of AI translation. By connecting the topic to both quality assessment of AI-powered translation tools and

human bias against such assessment, especially in the case of dialectical variations of SA, the results of this study highlight the level of development reached so far, and stress the necessity for such in-depth research and case studies to ensure the constant evolvement and reliability of AI translation and MT technologies.

4. References

- Abu-Rayyash, H. (2024). AI meets comedy: Viewers' reactions to GPT-4 generated humor translation. *Ampersand*, 12, 100162.
- Caliskan, A. (2023, August). Artificial intelligence, bias, and ethics. In *IJCAI* (pp. 7007-7013).
- Deokar, R., Nanjundan, P., & Mohanty, S. N. (2024, July). Transparency in translation: A deep dive into explainable AI techniques for bias mitigation. In *2024 Asia Pacific Conference on Innovation in Technology (APCIT)* (pp. 1-6). IEEE.
- Escudé Font, J. (2019). *Determining bias in machine translation with deep learning techniques* (Master's thesis, Universitat Politècnica de Catalunya).
- Fraser, K. C., Dawkins, H., & Kiritchenko, S. (2024). Detecting AI-generated text: Factors influencing detectability with current methods. *arXiv preprint arXiv:2406.15583*.
- Jakesch, M., Hancock, J. T., & Naaman, M. (2023). Human heuristics for AI-generated language are flawed. *Proceedings of the National Academy of Sciences*, 120(11), e2208839120.

- Ji, M., Bouillon, P., & Seligman, M. (2023). Cultural and Linguistic Bias of Neural Machine Translation Technology. *Studies in Natural Language Processing. Cambridge University Press*, 100-128.
- Li, X. (2025). Research on cultural subjectivity challenges and strategies for AI translation in the era of large language models. *English Language Teaching and Linguistics Studies*, 6, 114. <https://doi.org/10.22158/eltls.v6n6p114>
- Murayama, T., Miyazaki, K., Matsubara, Y., & Sakurai, Y. (2024). Linguistic landscape of generative AI perception: A global Twitter analysis across 14 languages. *arXiv preprint arXiv:2405.20037*.
- Nguyen, T. T., Hatua, A., & Sung, A. H. (2023, October). How to detect AI-generated texts? In *2023 IEEE 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)* (pp. 0464-0471). IEEE.
- Prates, M. O., Avelar, P. H., & Lamb, L. C. (2020). Assessing gender bias in machine translation: a case study with google translate. *Neural Computing and Applications*, 32, 6363-6381.
- Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). Can AI-generated text be reliably detected? *arXiv preprint arXiv:2303.11156*.
- Shahmerdanova, R. (2025). Artificial intelligence in translation: Challenges and opportunities. *Acta Globalis Humanitatis et Linguarum*, 2(1), 62-70.
- Wells, N. (2022). Translation as culture in the age of the machine. *Wasafiri*, 37(3), 77-80.
- Zaki, M. Z., Ahmed, U. (2024). Bridging linguistic divides: The impact of AI-powered translation systems on communication equity and inclusion. *Journal of Translation and Language Studies*, 5(2), 20-30.

Zbib, R., Malchiodi, E., Devlin, J., Stallard, D., Matsoukas, S., Schwartz, R., & Callison-Burch, C. (2012, June). Machine translation of Arabic dialects. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 49-59).

Zhu, T., Weissburg, I., Zhang, K., & Wang, W. Y. (2024). Human bias in the face of AI: The role of human judgment in AI-generated text evaluation. *arXiv preprint arXiv:2410.03723*.

5. Appendix: List of texts used along with their translations (across AI-powered translation tools) in no particular order

Text 1

E1-Ar	الدنيا قائمة قاعدة هون شو صاير
Literal Translation	The world's turned upside down here, what happened?
ChatGPT	The world's turned upside down here — what's going on?
DeepL	The world is standing still, here's what's going on
Google Translate	The world is standing here, what's happening?
DeepSeek	The world is just sitting here, what's happening?

Text 2

E2-Ar	عاملي فيها أبو العريف
-------	-----------------------

Literal Translation	He's acting like a know-it-all.
ChatGPT	He's playing the know-it-all with me
DeepL	A worker for me in Abu Al-Arif
Google Translate	He treated me like a chief
DeepSeek	He's acting all mysterious on me

Text 3

E3-Ar	طب يا مرسي تاني مرة لو سمحت ما تستعبطش
Literal Translation	Okay, Morsi. Next time please don't act like a fool.
ChatGPT	Okay, Morsi, next time, please don't play dumb.
DeepL	Morsi, please don't be a slave again.
Google Translate	Ok, Morsi, next time please don't make fun of me
DeepSeek	Okay, Morsi, next time, please don't act foolish.

Text 4

E4-Ar	هسه تريد فلوسك من النفط لو تخليها عمشاريع الدول
Literal Translation	Now do you want your money from oil, or would you rather leave it for state projects?

ChatGPT	Now you want your money from oil, or would you rather leave it for state projects?
DeepL	Do you want your money from oil if you give it to state projects?
Google Translate	Now you want your money from oil or leave it for state projects
DeepSeek	Now you want your money from oil, or would you rather leave it for state projects?

Text 5

E5-Ar	الألماس المخبري تشنه جنطة تقليد قاعدين نتحكم فتكوينها
Literal Translation	Lab-grown diamonds are just like knockoff bags as we control their composition
ChatGPT	Lab-grown diamonds are like a knockoff bag—we're the ones controlling their composition
DeepL	Laboratory diamonds are the most common types of diamonds that we can create.
Google Translate	Lab-grown diamonds are made by a traditional diamond-cutting machine that controls their formation.
DeepSeek	Lab-grown diamonds, what a knockoff bag! We're controlling their formation.

Text 6

E6-Ar	يعني اليوم شو مشكل صار بالشغل وقلن انا ماخصني. انا بشو وانتي بشو
Literal Translation	I mean what problem happened at work today! And I kept telling them it had nothing to do with me. Where am I and where are you?
ChatGPT	So what was the problem at work today? I told them, 'That's not my concern.' What does that have to do with me and you?
DeepL	I mean, today, what happened at work and they said, "It's not my business. I am what I am and you are what you are"
Google Translate	I mean, what happened today at work and they said it's none of my business. What's up with me and what's up with you?
DeepSeek	So today, what problem happened at work, and they said it's none of my business. What's it to me, and what's it to you?

Text 7

E7-Ar	طلعو اذكاء ستاتنا لما شنشلو حالن بالذهب بالزمنات
Literal Translation	Our grandmothers were quite smart when they draped themselves with gold back in the day.
ChatGPT	Our women were pretty smart back in the day when they decked themselves out in gold.
DeepL	The smartest women in the world are the ones who have never seen gold in their lives.

Google Translate	Our women are smart when it comes to wearing gold these days.
DeepSeek	Our women turned out to be smart when they invested in gold during those times.

Text 8

E8-Ar	سمبوسة ومسخن في اكلة واحدة يبقى اكيد اللي اخترعها فلسطيني لما فكر خارج الصحن
Literal Translation	Sambusa and Musakhan in one dish? It must have been a Palestinian who came up with this by thinking outside the ‘plate’!
ChatGPT	Samosa and Musakhan in one dish? Of course, the person who came up with it must be Palestinian—thinking outside the plate!
DeepL	It must have been invented by a Palestinian who didn't think outside the box.
Google Translate	Samosa and Musakhan in one meal, it must have been invented by a Palestinian who thought outside the plate
DeepSeek	Sambusa and musakhan in one dish? Definitely, the one who invented it was a Palestinian thinking outside the box

Text 9

E9-Ar	قلي حاضر بس ما طلعتش حاضر لأي حاجة
Literal Translation	He said okay, but he wasn't ready for anything.

ChatGPT	He told me 'okay,' but he wasn't actually ready for anything.
DeepL	Say yes, but never say yes to anything.
Google Translate	He said yes but he didn't say yes to anything
DeepSeek	He told me 'ready,' but it turned out he wasn't ready for anything

Text 10

E10-Ar	حمای لقی عندو ألف وون کوری حکالی بینفعوکی باشی؟ طلع کوزمو عامل أسبوع الأكل الكوري
Literal Translation	My father-in-law found 1000 Korean Won and asked if I could use them for anything. Apparently, Cosmo is having Korean food week.
ChatGPT	My father-in-law found 1,000 Korean won and asked me if I could use them for anything. Turns out, Cosmo is having a Korean food week.
DeepL	My father-in-law has a thousand Korean Won. What can I do for you? Cosmo's Korean Food Week.
Google Translate	My father-in-law found a thousand Korean won and asked me if they could help him. It turns out Cosmo is holding a Korean Food Week.
DeepSeek	My father-in-law found he had 1,000 Korean won and asked me, 'Can it get you anything?' Turns out Cosmo is doing Korean Food Week.