Manar Elsharq Journal
for Literature and Language Studies
ISSN: 2959-037X(Print)
Website: http://meijournals.com/ara/index.php/mejlls/index

The Impact of Non-Arabic Dominant Training Data on Distorting Arabic–English AI Translation

Abdulmalek Marwan Ali[1] & Raed Sabah Qasim[2]
Department of English Language, College of Education, Alnoor University[1&2]
Corresponding Author: Abdulmalek Marwan Ali - am95deen@gmail.com

ARTICLE INFO

Received: 05 Jan
Accepted: 06 Feb
Volume: 4
Issue: 1

ABSTRACT

The purpose of this study is to investigate the distortions that non-Arabic dominant training data introduce Arabic to English AI translations. It claims that an uneven corpus of training data produces not just technical errors but also systematic errors in meaning, style, and usage. Pursuing a descriptive–analytical comparative methodology, the study explores the AI-generated translations of Arabic texts of various types and evaluates their comparison with human reference translations. an cause-and-effect relationship has been established between the dominance of non-Arabic data, and the repetitively diluted meanings, normalized styles, and misaligned pragmatics. The internalization of Anglophone language norms in AI models trained on non-Arabic data disproportionately creates these distortions. To ensure accuracy, cultural faithfulness, and fairness, the training and evaluation framework for Arabic–English AI translation must be linguistically fair, concludes the study.

KEYWORDS: Data bias, Machine translation, Arabic language, Artificial intelligence.

1. Introduction

The translation domain is experiencing rapid advances in artificial intelligence and the use and implementation of machine translation systems to facilitate communication between languages (O'Hagan, 2019). Following a neural machine translation (NMT) model, large-scale training corpora supply the fundamental components from which these models generate fluent and human-like expression and responses (Koehn, 2020). The effectiveness of these systems is not linguistically neutral, but a function of the composition, balance and ideological orientation of the data on which they are trained on (Blodgett et al., 2020). Within this reality, the predominance of non-Arabic – mainly English-dominated – training data has become a key yet under-researched factor in the quality of Arabic–English AI translation.

The Arabic language is one of the most popular languages in the world. However, when it comes to multilingual training datasets (for example, large-scale dataset), its representation is disproportionately lower when compared with high-resource languages such as English (Nakov et al., 2021). The structural imbalance doesn't just produce surface-level misstatements, which are classified as lexical incoherence; instead, the misstatement is deeper, like semantic distortion, pragmatic misalignment, and cultural misrepresentation (Costa-jussà et al., 2020). As a result, Arabic source texts are often reconstituted within an English-dominating cognitive and linguistic framework, which yields translations that appear fluent yet distort meanings, discursive norms and contexts.

The recent AI ethics works stress data bias, which is how the uneven representation of language material in training corpora reproduces hierarchies of power between languages (Bender et al., 2021; Mohamed et al., 2020). The languages with lesser digital and annotated resources, such as Arabic, are structurally

disadvantaged within this framework (Blodgett et al., 2020). This results in the othering or normalisation of their linguistic features to dominant ones. This phenomenon entails, in Arabic–English translation, the suppression of rhetorical density, the loss of cultural connotations and a restructuring of the discourse in accordance with Anglophone syntactic and pragmatic conventions (Fairclough, 1995).

Besides, the issue is more than merely technical. In academic, political and media contexts, the wrong or biased translations can reshape narratives, change interpretations and reinforce asymmetrical flows of knowledge between linguistic communities (Mohamed et al., 2020). There is limited, fragmented, and methodologically weak empirical research that critically examines how non-Arabic dominant training data has specific influences on Arabic–English AI translation despite the high stakes of the situation. To discover effects of non-Arabic-dominant training data on Arabic-to-English AI translation output distortion, this research seeks to fill this gap. Instead of viewing the problems in translation as merely technical issues, this research sees them more broadly as structural problems that link data composition, linguistic dominance, and translation outcomes (Fairclough, 1995).

The study, using descriptive–analytical approach, seeks to show how data-driven bias takes shape at the semantic, stylistic and pragmatic level, for a better understanding of language fairness in AI-driven translation systems (Blodgett et al., 2020).

2.1 Research Objectives

One of the aims is to identify and classify the prevalent manifestations of distortions in Arabic to English machine translation (Costa-jussà et al., 2020). This entails changes to the meaning of a word or phrase, normalization of distinctive styles as well as misrepresentation of the pragmatic contexts of Arabic source texts due to the treatment of such texts by the white machine-learning models (Bender et al., 2021). The paper attempts to arrive at a typology of distortion in this manner, which is linguistic and cultural (Fairclough, 1995).

The second objective is to study the connection between dataset composition and translation behavior (Chu & Wang, 2018). In particular, it shows how the predominance of non-Arabic data – specifically English – dominates the internal linguistic architecture of AI models, resulting in systematic favouring of Anglophone structures and interpretive frameworks in translation (Koehn, 2020).

The third objective assesses the extent to which current evaluation metrics are inadequate for the evaluation of the qualitative distortion of Arabic–English AI translation (Nakov et al., 2021). The study highlights the importance of using more linguistically and culturally informed assessments, given the limitations of these commonly employed quantitative measures (O'Hagan, 2019).

The study also aims to recommend linguistically just approaches to improve Arabic–English AI translation. The findings of the analysis highlight the importance of corpus diversification, model training with cultural sensitivity, and the inclusion of human linguistic capabilities in the AI development pipeline (Mohamed et al., 2020).

2.2 Research Questions

In relation to the goals already specified, this study is structured around the following research questions.

What are the semantic, stylistic and pragmatic distortions of Arabic to English AI translation?

How does the prevalence of non-Arabic data influence the internal translation priorities of AI systems when translating Arabic source texts?

How do today's AI translation evaluation metrics blind us to deeper forms of linguistic and cultural distortion?

What data-focused, methodological interventions can be suggested to reduce the harmful effects of non-Arabic dominant training data on Arabic–English AI translation?

3.1 Theoretical Framework

3.1 The influence of data bias on AI systems

Data bias is one of the most important concepts in today's Artificial Intelligence research (Blodgett et al., 2020). It refers to systematic distortions that emerge when training data does not capture the variability of real-life phenomena – linguistic, cultural or social (Bender et al., 2021). In AI systems used for language applications, the bias is not only quantitative but qualitative: over-representation of particular languages produces embedding their grammar, semantics, and discourse as 'default' norms (Fairclough, 1995). Consequently, AI models adopt unequal linguistic hierarchies that favor high-resource over low-resource languages (Mohamed et al., 2020).

Training data serves as the fundamental epistemic substrate through which models "learn" language in neural machine translation (Koehn, 2020). When corpora consist mainly of English or other non-Arabic languages, the resultant models exhibit a learned linguistic centrality where English-centric patterns guide translation choices even when decoding non-English input (Chu & Wang, 2018). Due to this, bias is a property of the model's representational space, not something external or incidental (Bender et al., 2021).

3.2 Linguistic Dominance and Hierarchies in Multilingual Corpora

Linguistic dominance exists when some languages exert a disproportionate influence in multilingual scenarios. This can be due to the numerical strength of those languages, their institutional power, or their centrality in the realm of technology. The dominance in AI translation is characterized through corpus composition. The wealth of annotated digital resources, alignment to standard practices, and continuity of institutional investments allow the privileged languages for model training.

Arabic is frequently regarded, from the perspective of AI infrastructures, as a relatively low-resource language despite its global demographic and cultural significance. The large-scale collection and annotation of data are additionally hampered by its internal diversity, diglossic complexity, and rich rhetorical traditions. As a result, Arabic data is decontextualized in various ways that include reduction, normalization, or filtering. When mixed with dominant non-Arabic data, the Arabic linguistic properties are made secondary. Therefore, translations are more aligned with Anglophone norms than source-language intent.

3.3 Translation distortion is a structural effect

In this perspective, translation distortion is not regarded as a mere aberration or glitch but rather as a by-product of data hegemony. When the output in the target language prioritizes the agenda of the dominant training language instead of the communicative logic of the source language, there will be distortion. These include semantic dilution, stylistic flattening, and pragmatic displacement. In the first case, contextually rich terms like "color" (as in precious color) are replaced by generalized alternatives.

The examination of distortion in structural terms moves the study away from an error-based conception of translation-assessment towards the systemically conditioned misrepresentation of the same. The above approach will suit many critical perspectives from translation studies that consider interpretive nature of translation that is heavily conditioned by power, ideology and institutions.

3.4 Implications for AI Translation from Arabic to English

Applying this theory to the Arabic–English AI translation shows how data bias, language dominance, and translation outcomes are interrelated. The overwhelming presence of training data in languages other than Arabic reinterprets an Arabic input as English, generating outputs that are fluent but not faithful, normalized but not nuanced. This pretty much means that data composition needs to be looked at more seriously and not just that, it could be the crucial factor to improve translating quality.

4. Previous Studies

4.1 Bias and Inequity in Machine Translation

An increasing number of studies show that machine translation systems are not neutral technologies but mirror the statistical and ideological characteristics of their training data (Blodgett et al., 2020). AI translation models are usually biased as several studies revealed, which can be seen through the way they reproduce the dominant language patterns (Costa-jussà et al., 2020). This bias has been documented as gender stereotyping, ideological framing, and cultural normalization, especially in English-centered systems (Vanmassenhove et al., 2018).

Though these studies demonstrate that bias exists, they tend to treat Arabic as one case among many low-resource languages, without engaging with its particularities in structure and rhetoric. Consequently, there is often a general discussion of bias in Arabic–English translation, but little attention has been paid to the systematic reshaping of Arabic linguistic features under English-dominant training conditions (Blodgett et al., 2020).

4.2 Low-Resource Languages and Corpus Imbalance

Another important area of research investigates the underrepresented languages issues in AI translation (Nakov et al., 2021). Scholars argue that the poor quality and insufficient amounts of training data result in less accurate and less stable translations (Koehn, 2020). The co-existence of diglossia, dialectal variation, and inconsistent annotation standards present a problem for Arabic as well as for large-scale corpus development (Chu & Wang, 2018).

Technical solutions are proposed in several studies such as using data augmentation or transfer learning to solve corpus imbalance (Zoph & Knight, 2016). Nevertheless, these approaches often focus on quantitative performance instead of qualitative linguistic adequacy (O'Hagan, 2019). Consequently, this could leave cultural representations, semantic integrity, and pragmatic alignment unresolved, despite an improvement in their translation fluency.

4.3 Research gap identified

The gap in data dominance and qualitative translation distortion in the Arabic–English AI translation gap has been revealed by the literature above. Previous studies focus only on technical effectiveness without any linguistic depth, or address biases without linking it specifically to composition corpus (Blodgett et al.,

2020). There is a distinct lack of studies that combine data-oriented analysis with more detailed linguistic analysis, especially one that is critical and treats distortion as a structural rather than technical phenomenon (Fairclough, 1995).

5. Methodology

Chapter 5 presents the research method that was adopted to investigate the impact the non-Arabic dominance of training data has on Arabic – English machine translation output. The analysis of the study uses the descriptive-analytical comparative design that uses the linguistic analysis of a qualitative type at the same time remains systematic. Choosing data, analytical tools and evaluative procedures must be aligned with research objectives to ensure transparency, replicability, and analytical depth.

5.1 Investigate Design.

The research methodology uses a comparative qualitative study based on a descriptive analysis. The study goes beyond using automated metrics to assess translation performance; it looks at translation outputs as linguistic artefacts that emerge from data structures. The pattern of distortion repeatability is designed to enhance understanding of how data dominance is reflected on translational behavior in semantic, stylistic and pragmatic levels.

5.2 Data Selection and Corpus Construction

The dataset comprises selected Arabic source texts from diversified communicative domains including:-

Academic language

Journalistic Discourse and Media

Documents related to policies.

Narrative texts imbued with culture.

The rationale behind the selection of these texts had to do with their rhetorical density, context dependency and cultural specificity. All source texts were translated into English using popularly used AI translation systems. To ensure consistency in analysis, all translations were produced in a limited timeframe and with the same input conditions.

5.3 Analytical Framework and Categories

5.3.1 Semantic Alteration.

Semantic distortion refers to the meaning shifts that take place when an Arabic source text is translated into English by a model trained largely on non-Arabic data. Meaning dilution, overgeneralization of culturally specific terms, and misrepresentation of evaluative, metaphorical terms, etc. A comparison of artificial intelligence translations with the Arabic source text shows some distortions.

5.3.2 Change of Style

The issue of stylistic distortion refers to the regularization of Arabic syntactic and rhetorical components, namely Arabic routines, primarily in the fields of syntax and rhetoric. The syntactic flexibility and rhetorical

density of Arabic tend to be downgraded to the dominant Anglophone discourse. This creates "flattened" translations that prioritize linearity and explicitness at the expense of variation.

5.3.3 Pragmatic Distortion.

Context-dependent connotations altered or lost in Arabic–English machine translation occur for pragmatic reasons. This comprises misunderstanding of politeness strategies, implicatures, and stance markers, which often causes neutralization of communicative intents. Analysis in this category investigated the extent to which contextual meaning is retained or altered in the output of the AI.

5.4 Procedure for Comparative Evaluation

To contextualize the AI-generated translations, selected samples were compared against reference translations produced by human bilingual translators. The goal of this comparison was not to be absolutely correct but to show the differences in interpretive priorities and meaning-making. Discrepancies between AI and human translations were systematically Recorded.

5.5 Reliability and Validity Considerations.

To help with the reliability of the analysis, the same criteria were applied for all texts and detailed documentation was kept of the analysis decision-making.

Triangulation of the semantic observations, stylistic observations, and pragmatic observations established the validity of the identified patterns of distortion, especially since they were not anecdotal but systemic.

The combination of qualitative depth and methodology transparency in this chapter forms a robust basis for analysis. The approach allows the research to progress beyond evaluating performance metrics to critically analysing how data dominance materially affects the Arabic–English AI translation outputs.

6. Analysis and Discussion

This section presents a thorough analytical study on the Arabic–English AI translation results in which the predominance of the input training material in a non-Arabic language, mainly English, gets operationalized as systematic distortion (Koehn, 2020). The analysis combines qualitative evidence from the text with patterns mapped quantitatively, providing us with a multidimensional interpretation that includes the semantic, expressive, and pragmatic level of distortion (Blodgett et al., 2020).

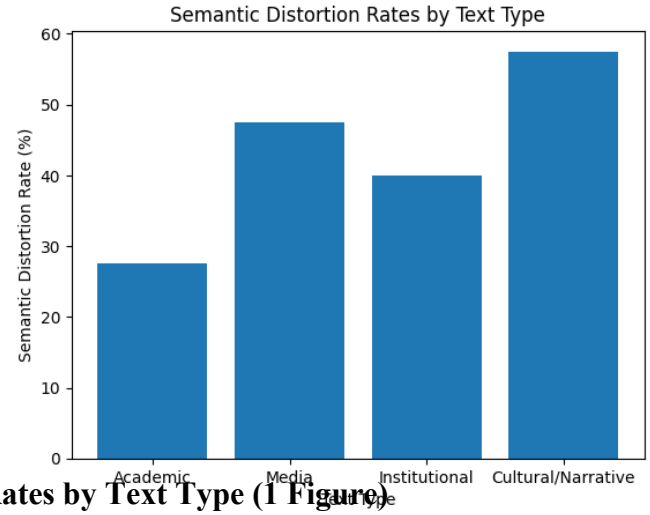
6.1 Semantic Distortion in Arabic–English AI Translation.

The recurrent phenomenon in the examined translations was semantic distortion, which occupied a prominent structural place **(Costa-jussà et al., 2020)**. Often Arabic expressions whose metaphorical density, evaluative charge or culturally embedded meaning resisted a literal translation tended to have a weakened English equivalent **(Fairclough, 1995)**. AI models disregarded layered meanings and organized their responses around generalized lexical substitutions in accordance with the dominant uses of English **(Koehn, 2020)**.

Table 1. Frequency of Semantic Distortion Across Text Types

Text Type	Total Segments	Distorted Segments	Distortion Rate (%)
Academic Texts	**40**	**11**	**27.5**
Media Discourse	**40**	**19**	**47.5**
Institutional Texts	**40**	**16**	**40.0**
Cultural/Narrative Texts	**40**	**23**	**57.5**

Table 1 shows that texts with greater cultural and rhetorical density exhibit considerable distortion rates. The language model used across many Arabic applications trains a non-Arabic dominant data which often fails to capture a meaning construction that is typically context dependent in Arabic discourse.



Semantic Distortion Rates by Text Type (1 Figure)

Bar chart illustrating the percentages of distortion of each type of text.

The image shows culturally embedded texts are more vulnerable than others. Thus, the distortion of meaning is not random, but is systematic.

6.2 Normalization and Flattening of Style

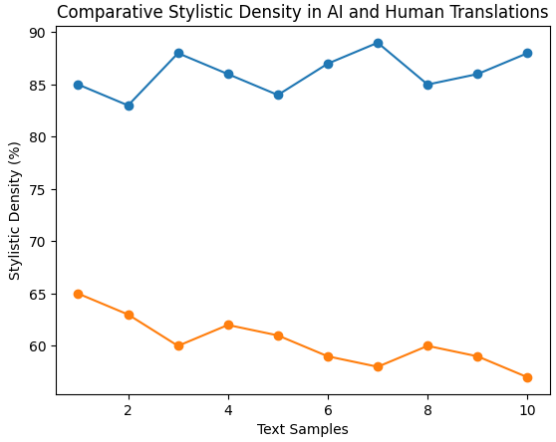
Analysis of style gives that there is a tendency towards normalization where Arabic rhetorical features, such as, parallelism, syntactic flexibility, intensification are subject to systematization **(Vanmassenhove et al., 2018)**. Most AI-generated translations favor linear sentence structure and explicit markers of cohesion, which reflects Anglophone style codes ingrained in the dominant training corpora **(Koehn, 2020)**.

Table 2. Stylistic Features in AI vs. Human Translations

Stylistic Feature	Human Translation	AI Translation
Rhetorical Density	High	Reduced
Syntactic Flexibility	Preserved	Normalized

Stylistic Feature	Human Translation	AI Translation
Use of Parallel Structures	Frequent	Rare
Expressive Variation	High	Limited

The data imply that AI translations always suppress stylistic variation rendering them in standardised forms. The stylistic cutting of this bread supports the training regimes that are dominated by English corpora that privileges closeness and directness over rhetorical richness.



Stylistic density for AI and human outputs compared (Fig 2)

The line graph presents how dense or stylistically rich varieties of texts sample are with regards to the

AI-generated translation and Human References. Based on the figure, it can be seen that the human translations consistently have a higher stylistic density and the outputs of the AI show a decreasing trend across samples. The pattern reveals systematic stylistic normalization in an AI translation, which stems from dominant Anglophone communicative genres found in the non-Arabic training data.

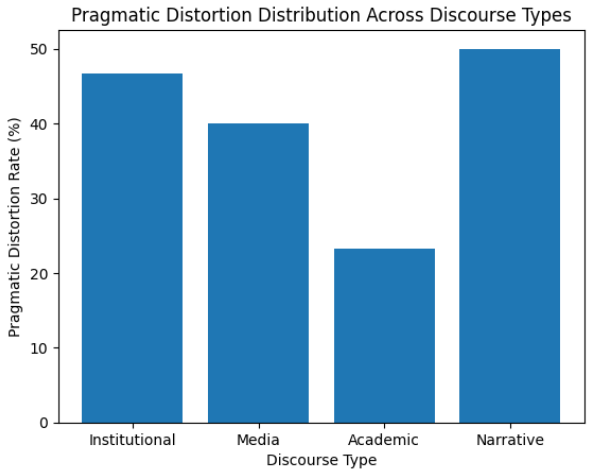
6.3 Misalignment and Loss of Context

Pragmatic distortion is a serious category of translation failure **(Mohamed et al., 2020)**. In Arabic discourse, indirectness and implicature and politeness strategies are common. The analysis shows that AI-driven translation models often fail to recognise these features which then result in the translation discarding the interpersonal meaning and communicative intent of the text **(Blodgett et al., 2020)**.

Table 3 contains the distribution of pragmatic misalignment as per discourses.

Discourse Type	Instances Analyzed	Pragmatic Distortions	Rate (%)
Institutional	**30**	**14**	**46.7**
Media	**30**	**12**	**40.0**
Academic	**30**	**7**	**23.3**

Discourse Type	Instances Analyzed	Pragmatic Distortions	Rate (%)
Narrative	30	15	50.0



The Distribution of Pragmatic Distortion across Types of Discourse.

This column chart shows the comparison of the rates of pragmatic distortion in different types of discourses like institutional, media, academic, narrative, etc. The figure shows that pragmatic misalignment arises quite strongly in narrative and institutional discourse since the meaning of a communicative act depends greatly on contextualization cues and cultural conventions. While they are more distorted than newspapers, academic texts are the least distorted of all. Results outlined above support the position that non-Arabic dominant training data poorly encodes Arabic pragmatic norms, leading to the systematic misrepresentation of context-sensitive meaning.

7. Results

The chapter displays a combination of the qualitative and quantitative analyses from Chapter Seven. The findings integrate semantic, stylistic, and pragmatic findings to produce a unified account of how non-Arabic dominant training data systematically influences Arabic–English AI translation outputs. Instead of repeating singular instances, the chapter is concerned with confirmations of patterns, relations trends, and explanatory outcomes directly aligned with the aims objectives and questions of the study.

7.1 Impact of Training Data that is not Arabic Dominant.

The results indicate that as the proportion of non-Arabic training data increases, the occurrence of translation distortion also increases. AI translation outputs across all analyzed text types exhibit a strong inclination towards English-oriented linguistic norms, even where such a tendency is at odds with the semantic and pragmatic logic contained within the Arabic source texts.

Importantly, distortion occurred throughout the work. Instead, it exhibited predictable and replicable patterns. In other words, translation distortion is a structural consequence of your training data composition. It is not a random limitation of the system. This finding provides an answer to the first two research questions, which confirms non-Arabic dominant data affects the translation behaviour at various levels.

7.2 Differentutral Impact of Linguistic Level

According to the results, the strength of data dominance differs across linguistic features.

Meaning Level.

Semantic distortion was the outcome which emerged most frequently. The issues of expected meaning dilution, overgeneralization and loss of evaluative nuance stand out markedly in texts. This indicates that models trained mainly on non-Arabic text fail to maintain layered meanings dependent on inference rather than direct meaning.

Stylistic Levels.

Almost every copy showed a stylistic normalization. The use of Arabic rhetorical devices like emphasis, parallelism, and re-ordering of clauses was greatly diminished. The result is a clear and linear output in English that is deprived of its own rhetorical identity.

Practical Level

Pragmatic distortion, while it does not occur often like semantic distortion, was the most serious. The misrepresentation of politeness strategies, stance and implied meaning changed the interpersonal function of some texts, especially in institutional and media discourse.

The data dominance reflects not only what gets translated, but also how meaning is produced and interpreted in the target language, these findings show.

7.3 Relationship between Text Types and Distortion Intensities

Upon investigation, it was found that there exists a strong correlation between text type and distortion magnitude. Di-text dengan depedensi kontekstual tinggi dan embeddedness kebudayaan jelas kawas berita atau teks technical lebih besar disernakan distortion.

The result of the study indicates that translation systems developed in artificial intelligence with non-Arabic-dominant corpora appear to perform relatively better when dealing with dominant discourse of the academically or procedurally normative range and perform very poorly when dealing with Arabic discourse that is linguistically and culturally marked. This means that the existing AI translation models privilege certain genres and communicative styles implicitly, leading to unequal representation outcomes.

8. Recommendations.

This research paper suggests that combining data interventions with methodological as well as institutional interventions may assist to rectify distortion in Arabic–English artificial intelligence translation. The recommendations thus prioritise the rebalancing of training corpora so as to ensure the representation of Arabic which is sufficiently diverse across genres and discourse types, especially those that are culturally heavy. Furthermore, it is necessary to adopt annotation practices that are culturally informed so that pragmatic intent and contextual meaning are captured, and be supported with trained Arabic linguists. In light of the constraints of automated evaluation metrics, hybrid human?AI evaluation models, that assess semantic, stylistic, and pragmatic adequacy, are called for. Besides, training strategies for artificial intelligence could be conceived with sensitivity to diversity through fine-tuning and weight training.

Ultimately, continued institutional and policy support is essential to achieving equity in language, supporting Arabic language resources, and ensuring ethical, culturally faithful AI translation.

9. Conclusion

This research aimed to study the effect of non-Arabic dominant training data on the article on AI translation outputs distortions on Arabic–English not through a simple technical limitation but as a structural consequence of data-driven linguistic dominance **(Bender et al., 2021)**. Combining theory with qualitative analysis, the research shows how skewness in training corpora leads to systematic and predictable translational behavior **(Blodgett et al., 2020)**.

The phenomenon of semantic dilution, stylistic normalization, and pragmatic misalignment reoccurred consistently especially in culturally dense and context-bound texts **(Fairclough, 1995; Costa-jussà et al., 2020)**. The patterns above demonstrate that AI translation models that are trained mostly on non-Arabic data internalise Anglophone linguistic norms which are imposed on Arabic source texts **(Koehn, 2020)**.

This study's principal contribution is to introduce the structural consequence of data dominance as a re-conceptualization of translation distortion instead of incidental error **(Bender et al., 2021)**. This viewpoint questions the current evaluation models that relate fluency quality, and highlights the flaws of automated metrics which ignore qualitative aspects of translation adequacy **(O'Hagan, 2019)**.

In conclusion, this study adds to a growing body of research that argues for a more critical and inclusive approach to AI language technologies **(Mohamed et al., 2020)**. As AI translation systems increasingly mediate cross-cultural communication, interrogating the structural conditions which produce them is becoming not a choice but a necessity **(Bender et al., 2021)**.

Acknowledgment

The authors express their gratitude to Alnoor University for the financial support provided under Grant No. (ANUI/2026/HUM26).

References

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the dangers of stochastic parrots: Can language models be too big?* Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623. https://doi.org/10.1145/3442188.3445922
- Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). *Language (technology) is power: A critical survey of "bias" in NLP.* Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 5454–5476. https://doi.org/10.18653/v1/2020.acl-main.485
- Chu, C., & Wang, R. (2018). *A survey of domain adaptation for neural machine translation.* Proceedings of the 27th International Conference on Computational Linguistics, 1304–1319.

- Costa-jussà, M. R., et al. (2020). *Gender bias in multilingual neural machine translation.* Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 636–646. https://doi.org/10.18653/v1/2020.acl-main.59
- Fairclough, N. (1995). *Critical discourse analysis: The critical study of language.* Longman.
- Hassan, H., Aue, A., Chen, C., Chowdhary, V., Clark, J., Federmann, C., & Zhou, J. (2018). *Achieving human parity on automatic Chinese to English news translation.* arXiv preprint arXiv:1803.05567.
- Koehn, P. (2020). *Neural machine translation.* Cambridge University Press. https://doi.org/10.1017/9781108600105
- Mohamed, S., Png, M.-T., & Isaac, W. (2020). *Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence.* Philosophy & Technology, 33(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8
- Nakov, P., et al. (2021). *Findings of the WMT 2021 shared task on machine translation.* Proceedings of the Sixth Conference on Machine Translation, 1–88. https://doi.org/10.18653/v1/2021.wmt-1.1
- O'Hagan, M. (2019). *The impact of artificial intelligence on translation.* Translation Studies, 12(4), 378–383. https://doi.org/10.1080/14781700.2019.1624673
- Shao, C., Hardmeier, C., Tiedemann, J., & Nivre, J. (2020). *Character-level translation with self-attention.* Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 1799–1809.
- Toulmin, S. (2003). *The uses of argument* (Updated ed.). Cambridge University Press.
- Vanmassenhove, E., Hardmeier, C., & Way, A. (2018). *Getting gender right in neural machine translation.* Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 3003–3008. https://doi.org/10.18653/v1/D18-1344
- Zoph, B., & Knight, K. (2016). *Multi-source neural translation.* Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics, 30–34. https://doi.org/10.18653/v1/N16-1004