A. Onofriichuk, *MA student*
S. Topachevskyi, *PhD in Phil., As. Prof.*
Zhytomyr Polytechnic State University

GENDER BIAS IN MACHINE TRANSLATION

This paper examines academic approaches to analyzing how gender bias manifests in machine translation (MT) systems and identifies unresolved issues within current research. Numerous MT tools, including Google Translate and DeepL, tend by default to masculine forms when translating gender-neutral sentences. For instance, the phrase "*They are a doctor*" is frequently rendered in Ukrainian as "*Він лікар*", despite the absence of gender specification in the source text.

Most researchers explain this by saying that MT systems learn from classical texts that already contain stereotypes [1, p. 75; 2, p. 96]. So they reproduce the portrayal of men as doctors and women as nurses. Some studies even suggest that human translators face the same issues, suggesting it's not just a machine problem; it's also a matter of how people think.

In this context, different solutions are proposed. One idea is to add gender hints to the input — like using "*he*" or "*she*" clearly [6, p. 85]. Another is to train the system on more balanced data, where women are depicted in strong roles and men in caring ones [1, p. 130]. Some researchers also say that this issue goes beyond grammar, as it concerns fairness and people's representation in language [1, p. 123; 3, p. 50].

Gender bias in MT is examined through outputs drawn from varied textual domains — such as media articles, administrative forms, and routine digital exchanges — where linguistic choices may implicitly reinforce gendered associations [2, p. 97; 4, p. 309].

But even with all this research, there are still gaps. Most studies focus on English and Western European languages [2, p. 96]. Ukrainian and other Slavic languages are rarely discussed, even though they also have gendered grammar and face similar problems. Additionally, most researchers examine large systems, such as Google Translate or DeepL. Smaller or local systems are often overlooked, although they are also used in real life [4, p. 309].

The frequent use of masculine forms in MT output stems not only from grammatical structure, but also from entrenched cultural norms that equate male-coded language with institutional authority and professional credibility [7, p. 430].

Another important question is: where does the bias come from? Researchers continue to debate whether bias stems primarily from system design or from the underlying training corpora. Others say it's the way the system is built [6, p. 85]. Marcello Prates et al. found that neural networks often fail to detect gender cues unless they are explicitly marked [5, p. 6365]. So fixing this means working on both the data and the model itself.

Different types of MT systems also handle gender differently. Adam Lopez explains how statistical systems rely on word frequency, which can reinforce stereotypes [10, p. 12]. Thierry Poibeau and Dorothy Kenny demonstrate that neural systems are better at fluency but remain weak in meaning and gender [7, p. 428; 8, p. 440]. Sneha Tripathi and Juran Krishna Sarkhel compare rule-based, statistical, and neural systems, indicating that each has its own strengths and weaknesses [9, p. 389].

Some researchers go even further, arguing that gender bias in translation extends beyond grammar. These systems are used in education, healthcare, and public services. If they constantly show men in powerful roles and women in passive ones, that can shape how people are seen — especially in official documents or public communication. That's why fairness in translation matters. It's not just a matter of getting the grammar right — it reflects a deeper need to represent people respectfully.

Ukrainian-speaking users may experience subtle emotional dissonance when encountering gendered MT output, particularly in contexts where translated language shapes self-representation — including CVs, evaluation forms, and online profiles [2, p. 97].

This broader view of gender bias in machine translation has led researchers to consider not only grammar, but also how biased translations affect people. As Dorothy Kenny observes, "*Translation is never neutral — it reflects the ideas in the data and the choices made by the people who build the systems*" [7, p. 430]. If MT tools keep linking certain jobs or roles to one gender, they help spread a quiet but steady form of misrepresentation. This becomes a real problem when translated texts are used in schools, government, or media, because they influence how people see gender roles.

Maria García González also points out that "*Automated translation systems are now used in many workplaces and international institutions*" [2, p. 97]. If these tools repeat stereotypes, they can hurt efforts to support gender equality. For example, a CV translated by a biased system might make a woman sound less confident or skilled, just because of word choice. These aren't just small mistakes — they can affect real opportunities.

To deal with this, some experts suggest making MT systems more open. Thierry Poibeau recommends that "*Developers should let users see how gender is handled and give them options to adjust it*" [6, p. 88]. This would give users more control and help avoid mistakes. But Sneha Tripathi and Juran Krishna Sarkhel explain that "*Rule-based systems are easier to control but less flexible, while neural systems sound more natural but are harder to guide*" [9, p. 391].

Another idea is to use training data that includes an equal number of male and female voices. Stefanie Ullmann says that "*Many MT systems are trained on texts that mostly show men, especially in science and business*" [1, p. 130]. If developers use more balanced data, the results can be more fair. But this takes time and care, because bias can come back through updates or user content.

This study explores the emergence of gender bias in MT systems. It proposes context-aware strategies for its mitigation, with specific attention to Ukrainian-language output and its impact on public-facing communication.

Finally, Anestis Karastergiou and Konstantinos Diamantopoulos remind us that "*Fixing gender bias in MT isn't just about changing code — it's about rethinking what translation should do*" [3, p. 62]. Should it focus on grammar, culture, or fairness? There's no single answer, but it's clear that gender bias in translation is a deep issue that connects languages, technology, and society. More research — especially in languages like Ukrainian — is needed to ensure that MT systems treat all people fairly.

Gender asymmetries in MT output reveal systemic imbalances in how language technologies encode and circulate social roles. Remediation efforts must extend beyond algorithmic refinement to include sustained attention to inclusive representation in digital translation practices.

REFERENCES

1. Ullmann S. Gender bias in machine translation systems // Artificial Intelligence and Its Discontents. – Cham : Springer, 2022. – C. 123–144.
2. González M. G. Gender bias in translation automation // The Social Impact of Automating Translation. – London : Routledge, 2024. – C. 96–122.
3. Karastergiou A. P., Diamantopoulos K. Gender Issues in Machine Translation // Transcultural Journal of Humanities and Social Sciences. – 2024. – T. 5(1). – C. 48–64.
4. Danylov H., Balakirieva V., Vasylenko K. Machine Translation Systems and Their Specifications // Lingstud. – 2021. – №33. – C. 309–315.
5. Prates M., Avelar P. H., Lamb L. C. Assessing gender bias in machine translation: A case study with Google Translate // Neural Computing and Applications. – 2020. – T. 32(10). – C. 6363–6381.
6. Poibeau T. Machine Translation. – Cambridge : MIT Press, 2017. – C. 85–92.
7. Kenny D. Machine Translation // The Routledge Handbook of Translation and Philosophy. – London : Routledge, 2018. – C. 428–445.
8. Rizzardi S. Gender bias in machine translation: Causes, characteristics, and effects. – Padua : University of Padua, 2023. – C. 440–455.
9. Tripathi S., Sarkhel J. K. Approaches to machine translation // Annals of Library and Information Studies. – 2010. – T. 57. – C. 388–393.
10. Lopez A. Statistical machine translation // ACM Computing Surveys. – 2008. – T. 40(3). – C. 1–49.