

# Evaluation of adaptive machine translation from a gender-neutral language perspective

Aida Kostikova,  
Dr Joke Daems and Dr Todor Lazarov



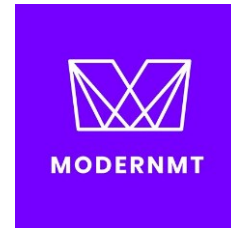
# What is adaptive machine translation?

- Adaptive MT learns from the edits on-the-fly. When a translator confirms a segment, an engine processes the source/target pair as correct and uses it as an example to provide future translation suggestions

- baseline model



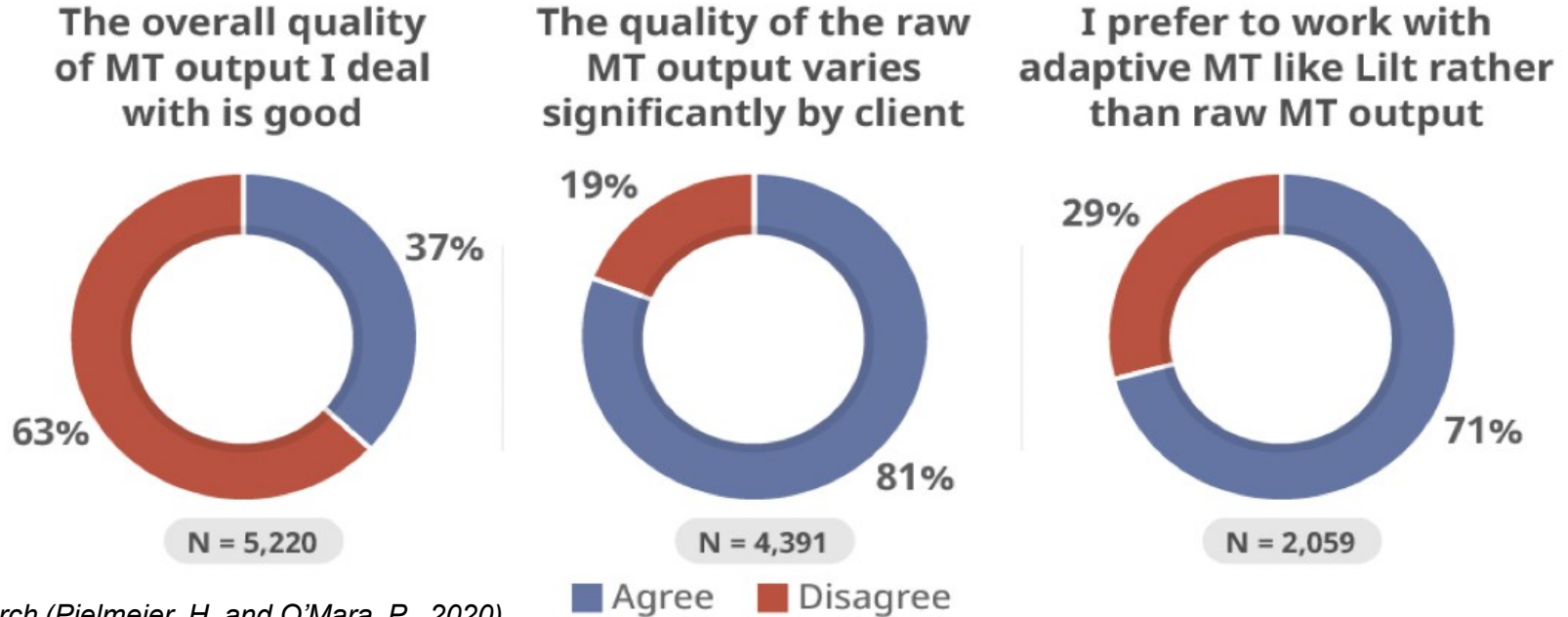
- adaptive model



# What is adaptive machine translation?

- Adaptive MT learns from the edits on-the-fly. When a translator confirms a segment, an engine processes the source/target pair as correct and uses it as an example to provide future translation suggestions

## Linguist Perspectives on Machine Translation



# 1. Background

## 1.1. Research problems and objectives

To which extent can adaptive machine translation engines adjust to GNL use?



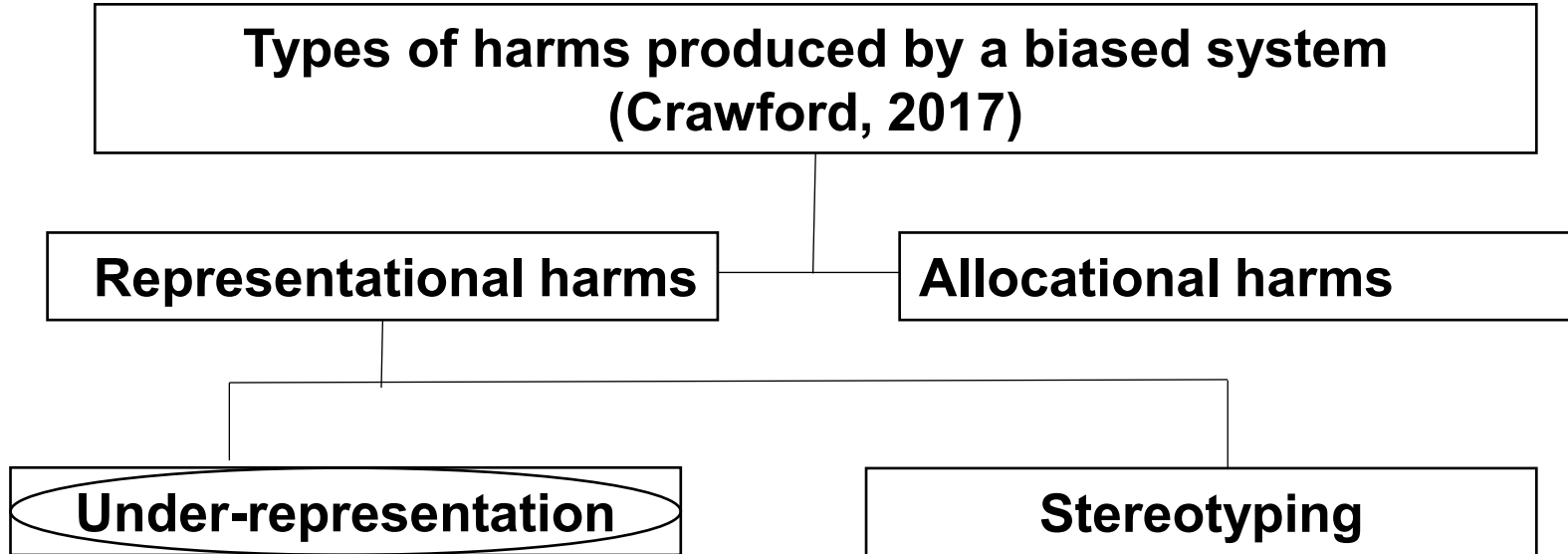
**Goal 1:** assess the baseline of an adaptive MT engine from the standpoint of misgendering



**Goal 2:** assess the speed of adaptation of the engine in attempting to ensure gender-neutrality in translation

## 1.2. Gender bias in MT

Biased system is an MT model that *systematically* and *unfairly* discriminates against certain individuals or groups in favor of others (Friedman and Nissenbaum, 1996)



## 1.2. Gender bias in MT

### **Mitigating strategies:**

- Model debiasing (gender tags, debiased word embeddings, lattice re-scoring modules) (Vanmassenhove et al., 2018)
  - Evaluation methods (Rudinger et al. 2018; Zhao et al. 2018)
  - Working with training data (Zhao et al. 2017)
  - follow a more focused approach within NLP (Savoldi et al., 2021)
- BUT they:**
- lack human-computer interaction component
  - are reduced to the binary dichotomy (Dev et al., 2021)

## 1.3. Gender-neutral language: overview

### Language type

(Stahlberg et al., 2007):  
Notional (English),  
grammatical gender  
(German) and genderless  
languages (Turkish)

### Gender system:

Binary or non-binary?

### Gender-fair approach:

Neutralization (no mention  
of gender) or inclusion  
(explicitly challenges  
binary notions of gender)  
del Río González, 2021

### Gendersternchen (gender asterisk)

Nouns	Personal	Possessives	Articles	Question Pronouns
Student*in Student*innen	Sie*er	Ihre*seine	der*die	Welche*r?

### De-E-System<sup>1</sup>

Case	Def. Articles.	Indef. Articles.
Nom.	de gute Lehrere	ein gute Lehrere
Dat.	dern guten Lehrere	einern guten Lehrere
Acc.	de gute Lehrere	ein gute Lehrere

1. Verein für geschlechtneutrales Deutsch) (<https://geschlechtsneutral.net>)



# 2. Experiment

## 2.1. Setup

### Material:

- Rulebook of the International Quidditch Association (English)
- 45 segments (average length: 65 characters)

### Conditions:

- English-German language pair
- The overall quality was ignored, only “problematic” units were post-edited

### Metrics:

- KSR or keystroke ratio: the number of keystrokes divided by the total number of characters in the reference (Daems and Macken, 2019)
- CharacTER (Wang et al., 2016), a character-based variant of TER

## 2.2. Results

- English gender-neutral pronouns are translated incorrectly

The options available to that player are determined by where they would resume by default in the absence of this rule as follows:

Die Optionen, die diesem Spieler zur Verfügung stehen, werden davon bestimmt, wo sie in Ermangelung dieser Regel standardmäßig fortgesetzt würden, wie folgt:

- Gender-neutral **lexical** choices from the first segment

A team may not have more than four players who identify as the same gender in play at the same time.

Ein Team darf nicht mehr als vier Spieler haben, die sich gleichzeitig als gleichgeschlechtlich identifizieren.

- Generic masculine:

The speaking captain has the power to speak for the team when conversing with officials.

Der sprechende Kapitän hat die Macht, für das Team zu sprechen, wenn er mit Beamten spricht.

# The goal is to reach KSR and CharacTER with the 0 value:

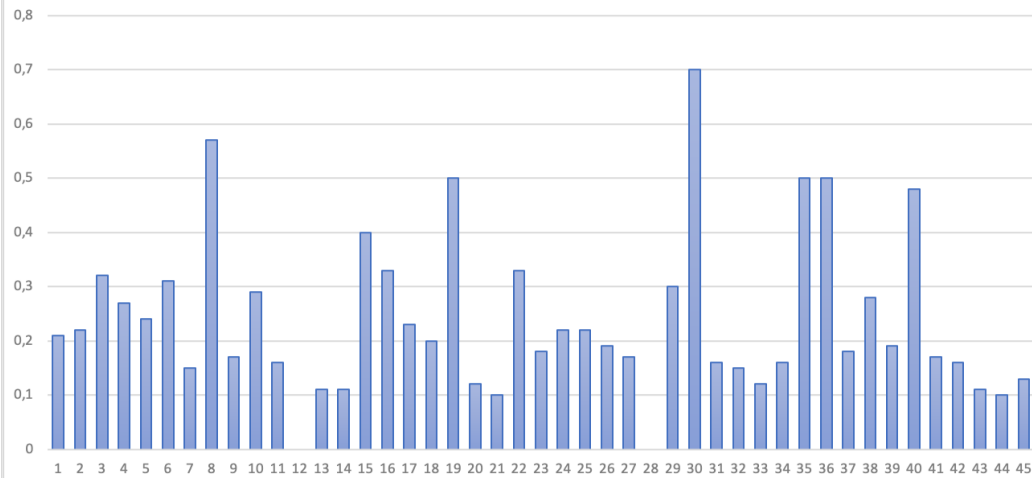
<b>Source</b>	Officials may direct any person to stop talking with any official.
<b>Suggestion</b>	
<b><u>Translation</u></b> 	Die Beamterne können jede Person anweisen, das Gespräch mit einerm Beamtere zu beenden.



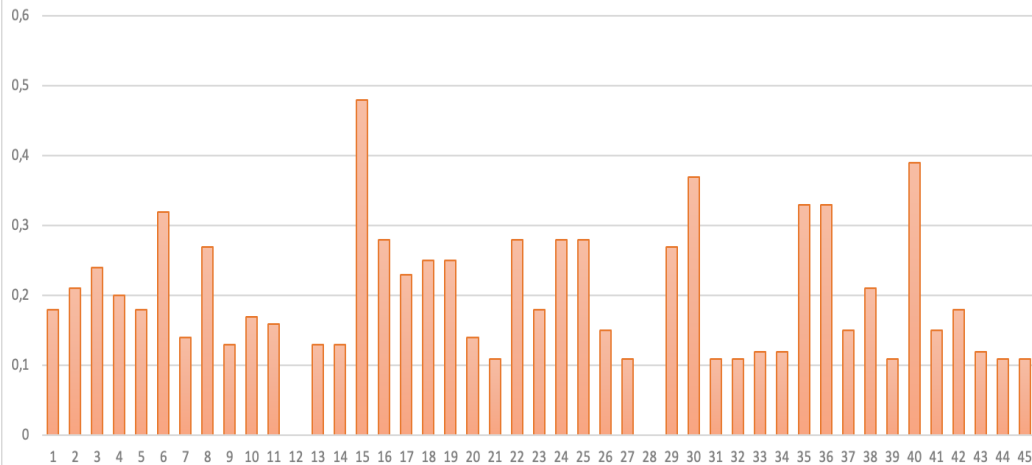
<b>Source</b>	Officials may direct any person to stop talking with any official.
<b>Suggestion</b>	
<b><u>Translation</u></b> 	Die Beamter*innen können jede Person anweisen, das Gespräch mit eine*m Beamter*in zu beenden.



De-E-System: KSR



De-E-System: CharacTER



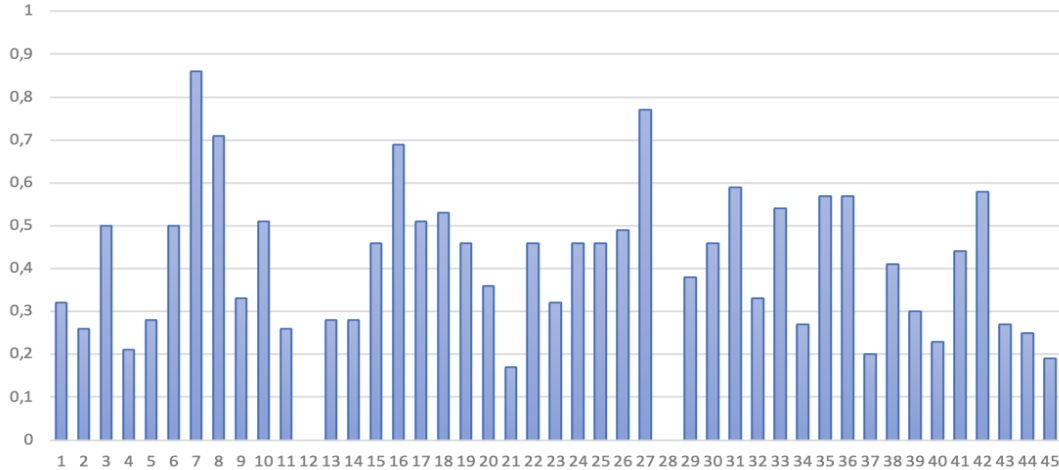
## De-E-System results:

- No improvement in the adaptation speed (just a few segments with the KSR and CharacTER with the 0 values)
- Some improvements (segments 12 and 28) are due to the translation memory matches

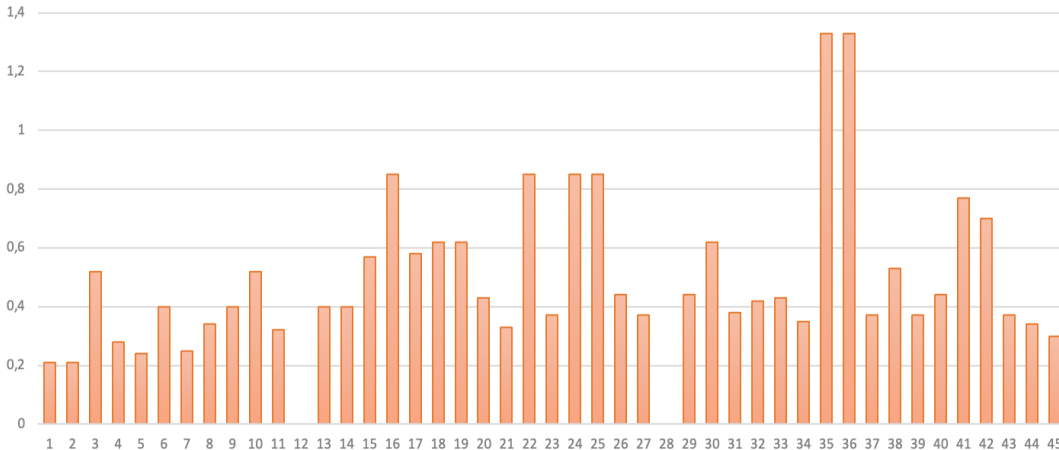
## Gender asterisk results:

- No improvement in the adaptation speed (just a few segments with the KSR and CharacTER with the 0 values)
- Some improvements (segments 12 and 28) are due to the translation memory matches

Gender asterisk: KSR



Gender asterisk: CharacTER



## 2.2. Results

### Other observations:

- Typographic characters did not have any influence on the rest of the sentence

Auswechselfpieler\*innen sind nicht verpflichtet,



Symbol mismatch

- A very common error with a symbol mismatch
- TM match was suggested with the both of the strategies

# 3. Conclusions



### 3. Conclusions

1. No progress in adaptation speed, except for the cases of TM auto-propagation;
2. Bias present in the initial output is linked to the grammatical gender, not to lexical invariability;

i. Wenn der\*die ursprüngliche sprechende

---

i. Wenn der\*die ursprüngliche sprechende **Kapitän\*in** des Teams für das Spiel legal auf die Bank oder den Pitch zurückkehrt, werden sie die Rolle des sprechenden Kapitäns wieder aufnehmen.

# References

- Crawford, K. (2017). The trouble with bias. In Conference on Neural Information Processing Systems (NIPS) – Keynote.
- Daems, J., & Macken, L. (2019). Interactive adaptive SMT versus interactive adaptive NMT: a user experience evaluation. *Machine Translation*, 33(1), 117-134.
- Dev, S., Monajatipoor, M., Ovalle, A., Subramonian, A., Phillips, J. M., & Chang, K. W. (2021). Harms of gender exclusivity and challenges in non-binary representation in language technologies. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp. 1968–1994
- Friedman, B., and Nissenbaum, H. (1996). Bias in computer systems. In *ACM Transactions on Information Systems (TOIS)*, 14(3), pp. 330–347.
- María del Río-González, A., (2021). To Latinx or not to Latinx: a question of gender inclusivity versus gender neutrality. *American Journal of Public Health*, 111(6), pp.1018-1021.
- Pielmeier, H. and O'Mara, P. (2020). The state of the linguist supply chain: Translators and interpreters in 2020 (Report). Boston: CSA Research
- Rudinger, R., Naradowsky, J., Leonard, B., Van Durme, B. (2018) Gender bias in coreference resolution. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2 (Short Papers), pp. 8–14.
- Savoldi, B., Gaido, M., Bentivogli, L., Negri, M., Turchi, M. (2021) Gender Bias in Machine Translation. In *Transactions of the Association for Computational Linguistics*; 9, pp. 845–874.
- Stahlberg, D., Braun, F., Irmen, L., & Sczesny, S. (2007). Representation of the sexes in language. In *Social communication*, pp. 163-187.
- Vanmassenhove, E., Hardmeier, C., & Way, A. (2018). Getting gender right in neural machine translation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 3003–3008
- Wang W, Peter JT, Rosendahl H, Ney H (2016) CharacTer: translation edit rate on character level. In: *Proceedings of the first conference on machine translation: volume 2, shared task papers*. Association for Computational Linguistics
- Zhao, J., Wang, T., Yatskar, M., Ordonez, V., & Chang, K. W. (2018). Gender bias in coreference resolution: Evaluation and debiasing methods. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2 (Short Papers), pp. 15–20.