

Introducing Fairslator: a machine translation bias removal tool

Michal Měchura

Our next 15 minutes together...

On-screen demo of Fairslator

•www.fairslator.com

Motivation for Fairslator

•Bias, ambiguity and all that

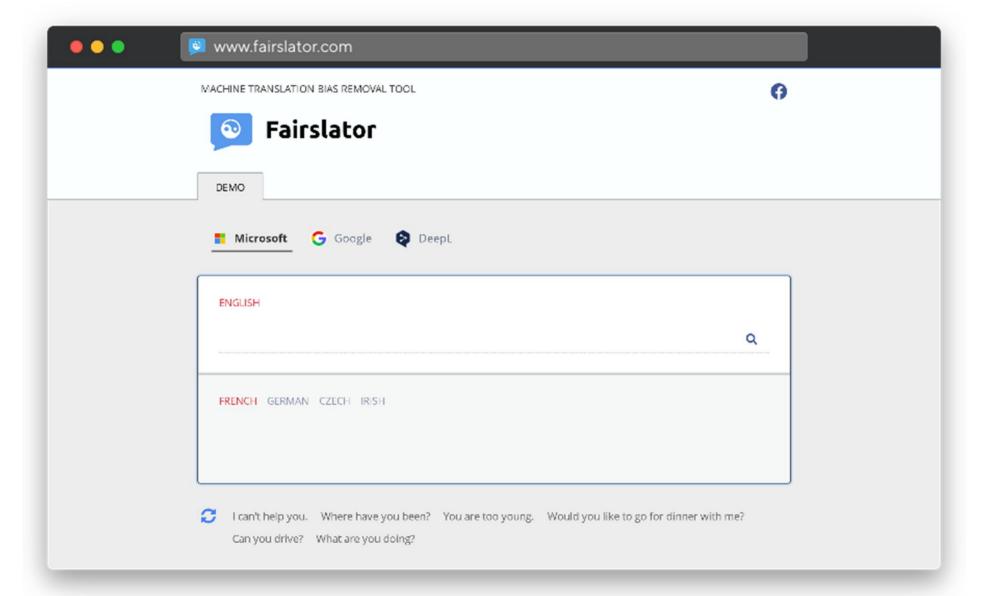
Technology behind Fairslator

- .Human in the loop
- Rule-based re-inflection

Future plans for Fairslator

Similar work elsewhere

If time allows



English

I am a good student.



French

Je suis un bon étudiant. Je suis une bonne étudiante.

German

Ich bin ein guter Student. Ich bin eine gute Studentin.

Czech

Já jsem dobrý student.

Já jsem dobrá studentka.

Russian

Я хороший студент.

Я хорошая студентка.



ENGLISH I am a good student. FRENCH GERMAN CZECH IRISH



ENGLISH I am a good student. FRENCH GERMAN CZECH IRISH







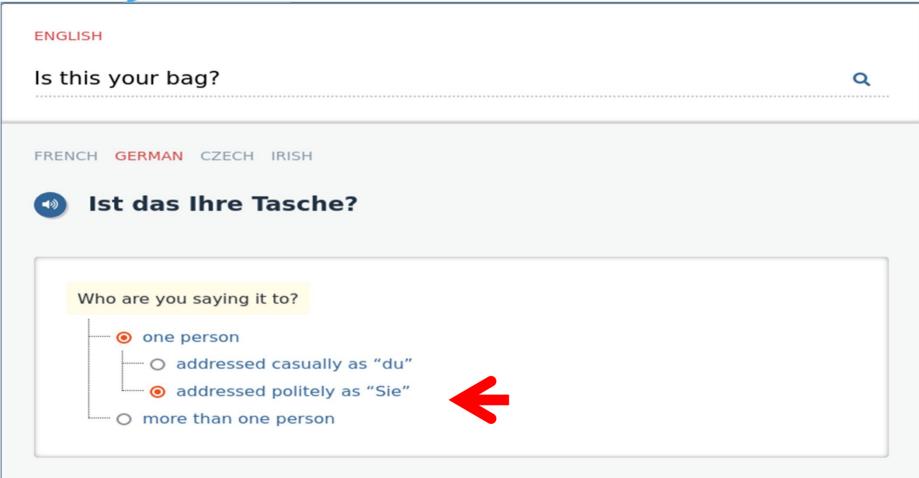




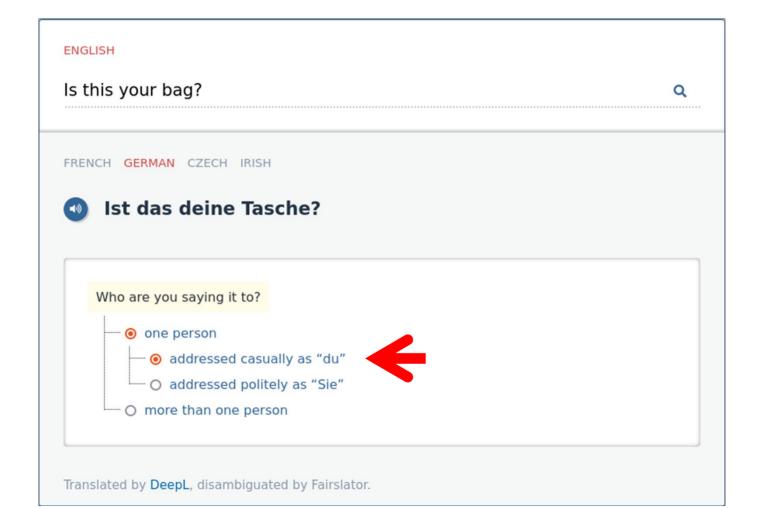




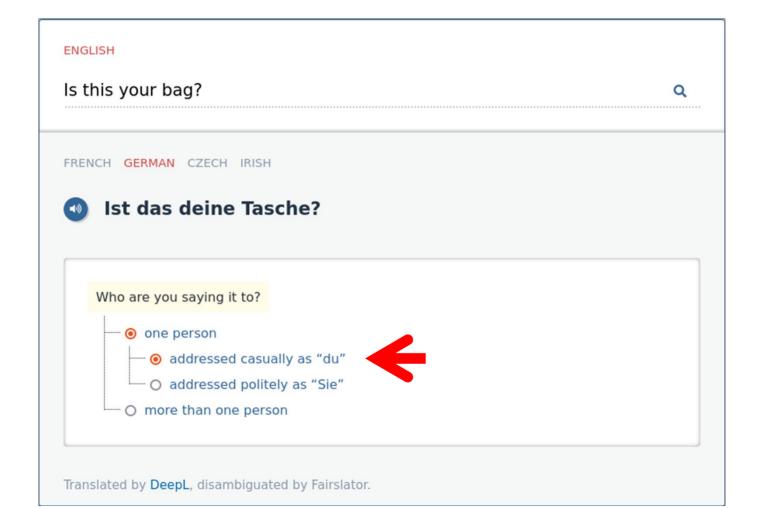
Translated by DeepL, disambiguated by Fairslator.



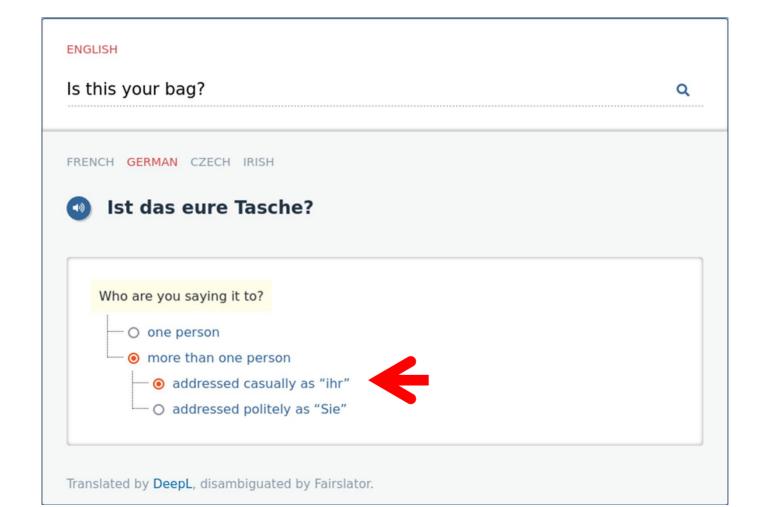




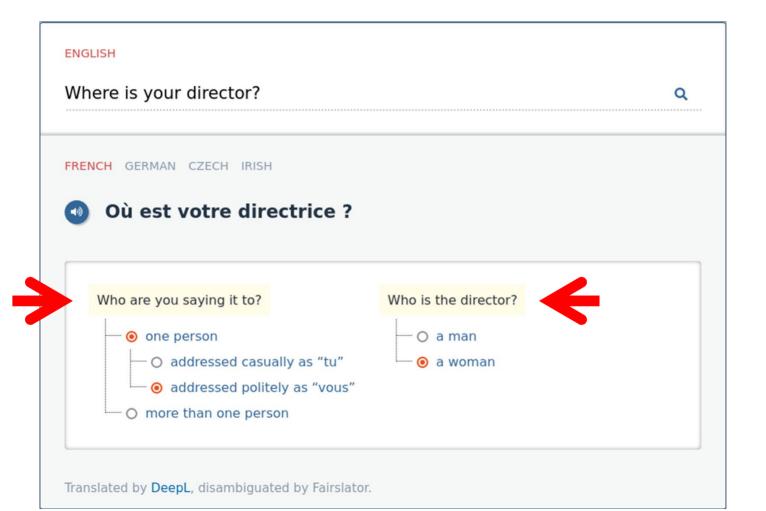




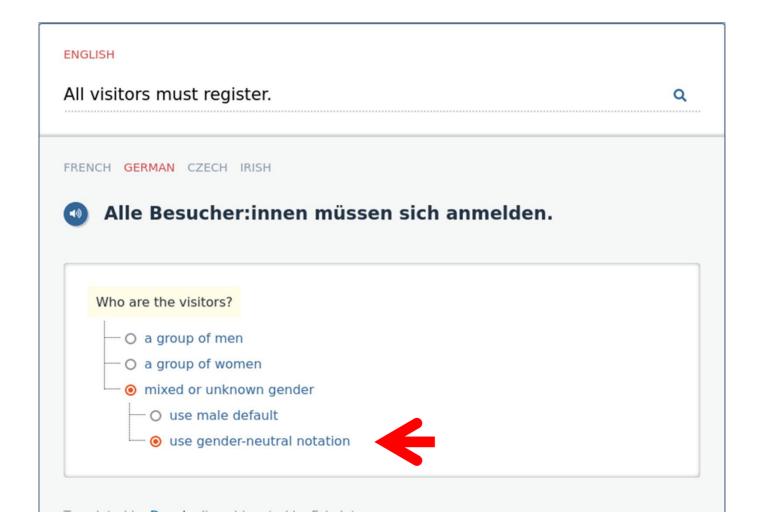


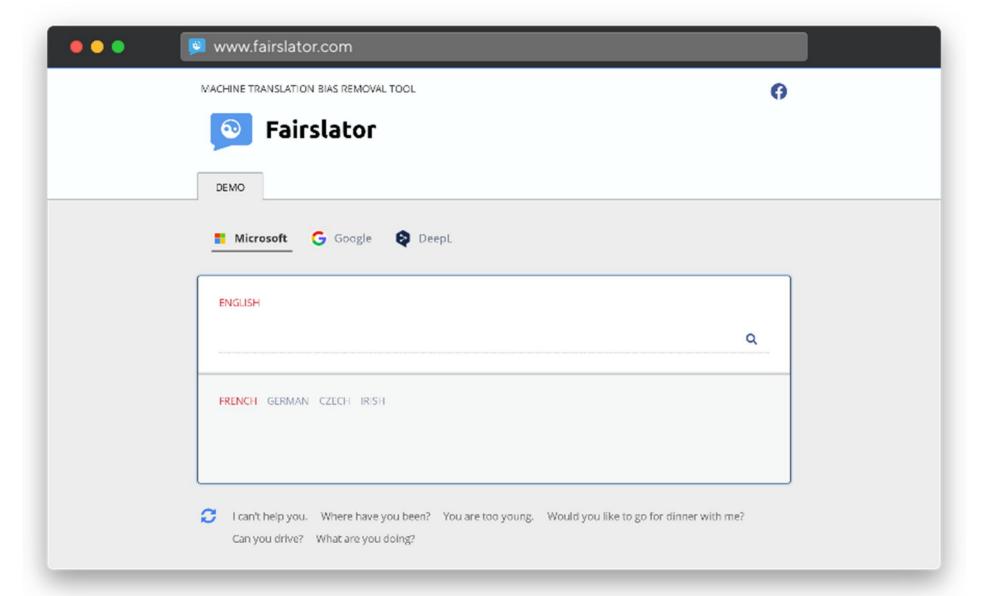












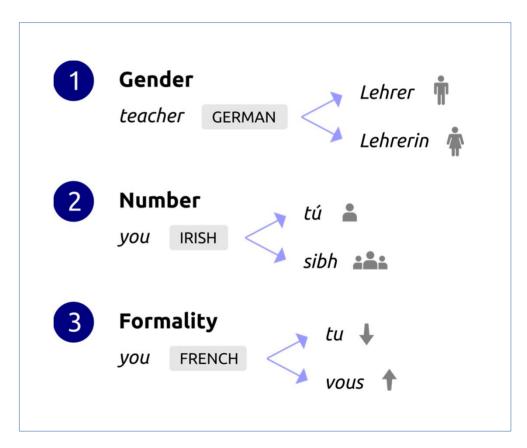
Machine translation bias

A side effect of machine learning

- •Biased training data → biased language models
- Overgeneralisation

Impact of biased MT

- Incorrect translations
- .Perpetuation of stereotypes



You don't hate bias, you hate ambiguity

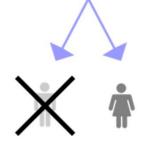
"Languages differ essentially in what they <u>must</u> convey and not in what they <u>may</u> convey."

— Roman Jakobson, 1959

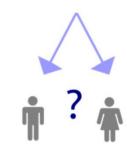


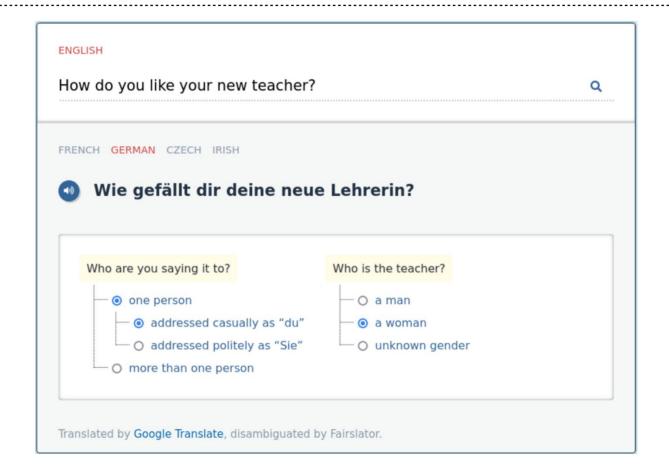
Resolvable versus unresolvable ambiguities

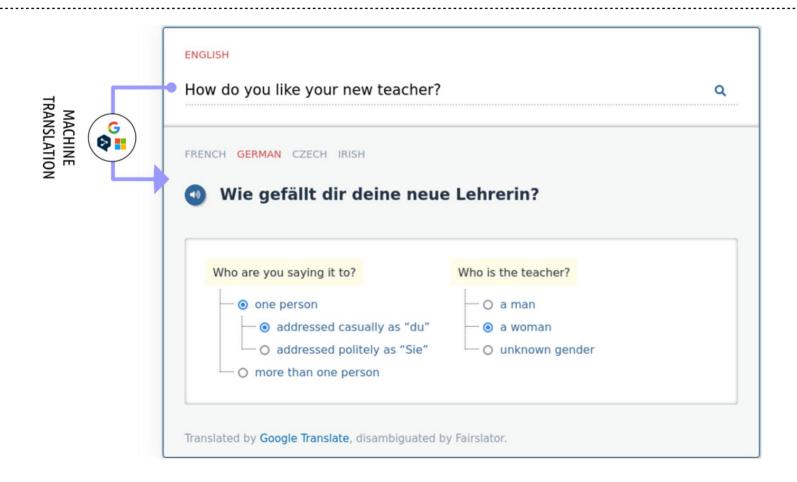
She is a teacher.

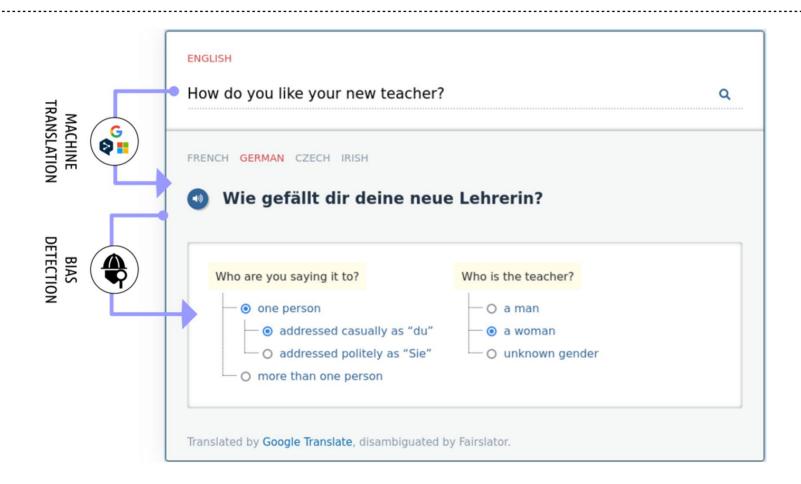


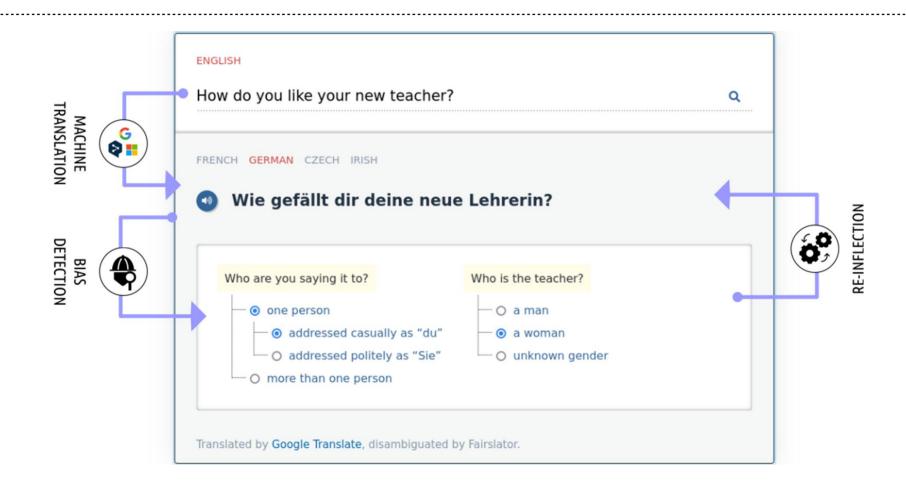
I am a teacher.











Fairslator is...

A plug-in for existing MT systems

- .Treats MT as a black box
- Only examines its output

A set of rule-based algorithms

- No massive machine-learning infrastructure needed
- •Workload: 3 5 person-days per language pair

I am very happy. **FRENCH** Je suis très heureux. Je suis très heureuse. Listener axis Where have you been? GERMAN Wo sind Sie gewesen? Wo bist du gewesen? Wo seid ihr gewesen? Bystander axis This is my new doctor. CZECH

Speaker axis

Měchura, M. (2022) 'A taxonomy of bias-causing ambiguities in machine translation of this Workshop on Gend



Thank you.

For more a deeper dive, read the Fairslator Whitepaper





www.fairslator.com