

# LA RÉAPPARITION

Automatic Translation from Standard French to Inclusive French Based on Syntactic Criteria

Léo Mergret  
leo.mergret@etu.u-paris.fr

Alice Dagand-Drillat  
alice.dagand@etu.u-paris.fr



## - INTRODUCTION -

Often misinterpreted, known to be more difficult to read and write, perceived as a “mortal peril” by some, hill to die on to some; no one will have missed it, inclusive writing is at the very core of many controversies since it has been created. Rather than fueling the endless debate about its relevance, and as part of our IDEEL introduction to research course within the University of Paris, we have chosen to try and set up an automatic translator, converting standard French to inclusive French. For matters of time and feasibility, we focused on changes that depend on morpho-syntactic criteria.

As far as we know, there is currently no achieved project on the subject of automatic translation to inclusive French, although associations are working on it and already published inclusive lexicons [3].

## - STANDARDS USED -

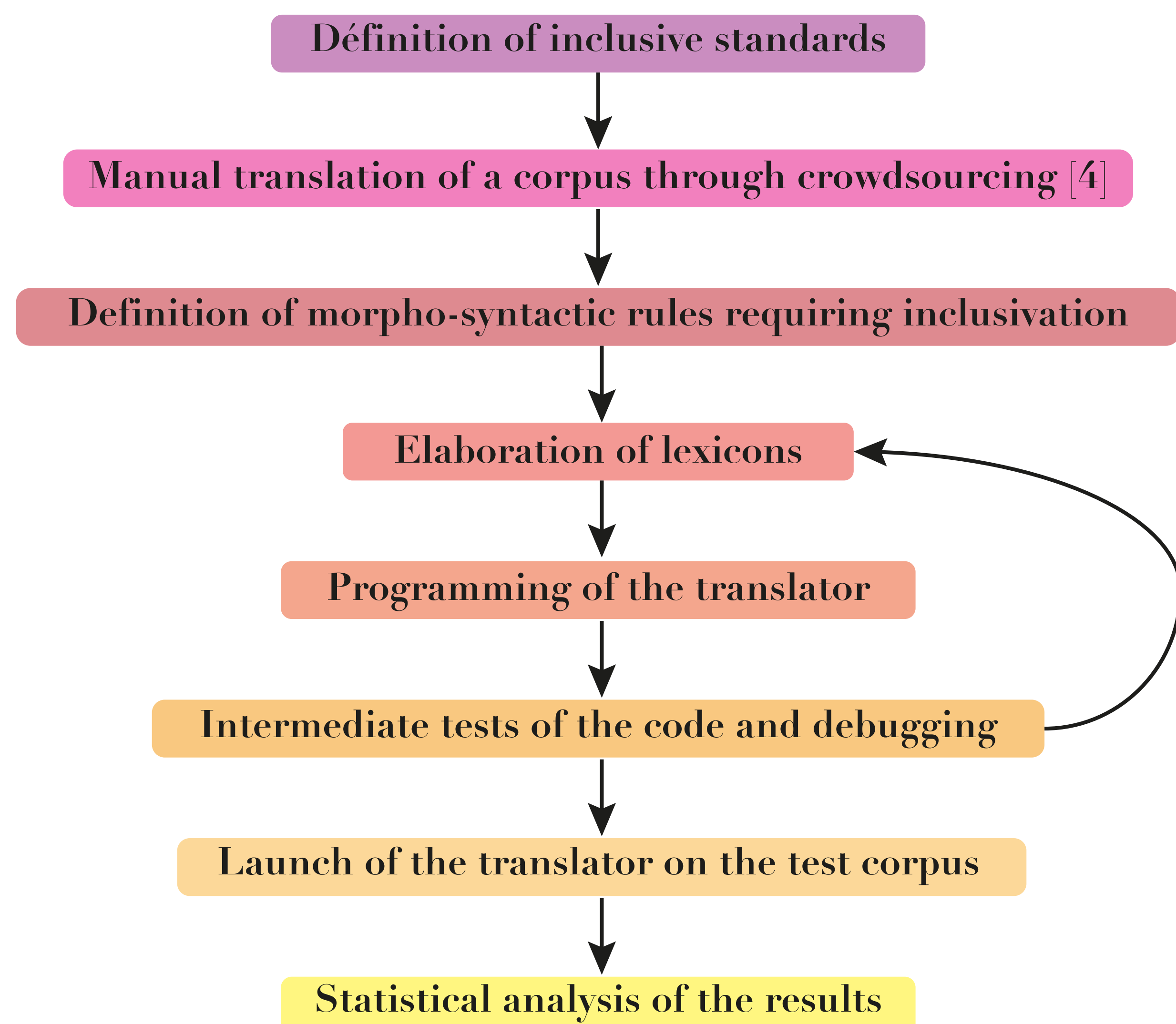
There are several “levels” of inclusive writing, centred mostly on the use of the midpoint [2]. We chose to apply a “moderate” level, for the sake of reading fluidity and accessibility.

Inclusive writing is a recent phenomenon[1], for which universal standards are yet to be established; we have then had to make a choice regarding the standards we followed for this project.

Here are the rules we followed:

- Use of the midpoint only for words referring to humans, and only when the masculine form is strictly included in the feminine form of the word;
- Proximity agreement;
- Feminisation of professions and titles;
- Substitution of gendered terms for gender-neutral terms whenever possible.

## - METHOD -



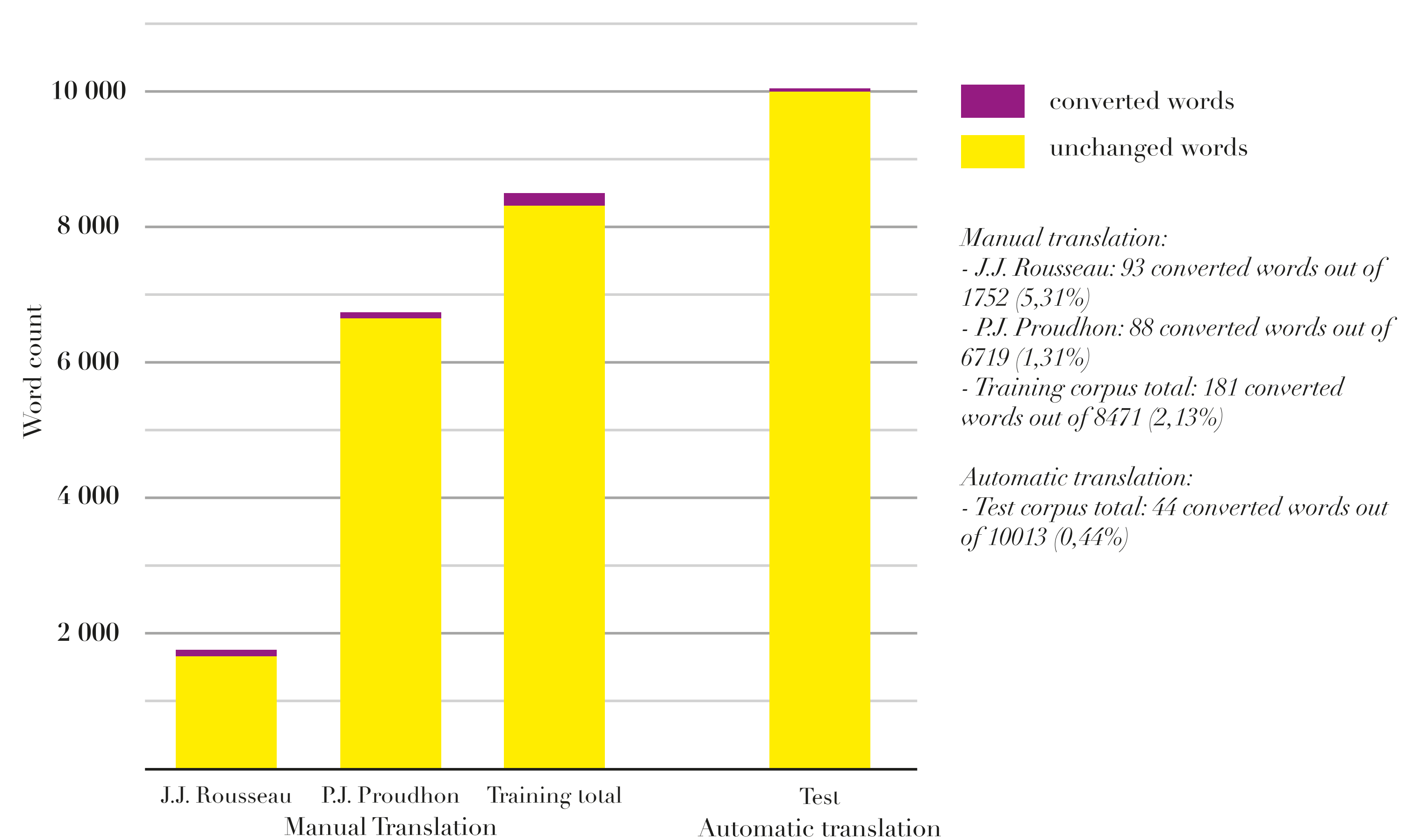
Our translator is made of:

- 3 lexicons (professions, nouns referring to mixt human groups, nationalities) in .txt format, containing words in the masculine form and their feminine/gender-neutral equivalent.
- The text we want to translate [5], in .txt format; it must have been previously PoS-tagged, and show, for each word, at least: the in-context version, the uninflected version, the syntactic category, the gender, and the number (if applicable).
- A programm in Python. It loads all 3 lexicons and the untranslated text, then applies all conversion rules and produces a document in .txt format, containing the translated text (without tags). It also show details regarding the changes implemented.

## - ABSTRACT -

At a time where we know to what extent language influences thought, inclusive language – which aims at reducing gender inequalities – is on the rise. Is it possible to create an automatic translator from standard French to inclusive French? In order to do this, we set up an algorithm based on explicit rules. We chose to focus on situations where neutral masculine, used instead of an inclusive form, can be detected through morpho-syntactic criteria. This program has been trained with a corpus of two essays in both versions: standard French, and inclusive French. It has then been tested on a PoS-tagged corpus in standard French.

## - RESULTS -



Proportion of Converted Words in Manually and Automatically Translated Corpora



Syntactic Categories of Converted Words in Automatically Translated Corpus

Manual translation shows a significantly higher amount of changes than automatic translation. However, considering the size of reviewed data and the variation within manual texts, this result should be taken with a grain of salt.

Obviously, focusing only on changes implementable via syntactic criteria avoids a significant amount of words requiring inclusivation, as we necessarily miss all semantic and pragmatic hints.

Moreover, in order not to overly complexify the code, changes in agreement are only done locally: if a masculine noun is replaced with its gender-neutral form, any adjective referring to it cannot be detected further than 2 words before or after.

## - BIBLIOGRAPHY -

1. <https://www.relire-et-corriger.net/ecriture-inclusive>
2. <https://lessalopettes.wordpress.com/2017/09/27/petit-guide-pratique-de-lecriture-inclusive/>
3. <https://inluzore.fr>
4. <https://www.gutenberg.org/>
5. [https://github.com/UniversalDependencies/UD\\_French-GSD](https://github.com/UniversalDependencies/UD_French-GSD)  
"Qu'Est-ce que la propriété ?" (Pierre-Joseph Proudhon)  
"Du Contrat social" (Jean-Jacques Rousseau)