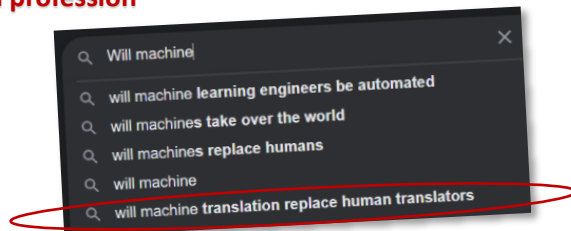
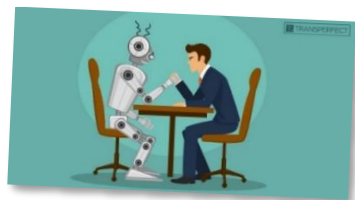


Discourse markers, a bastion of human translation against machine translation?

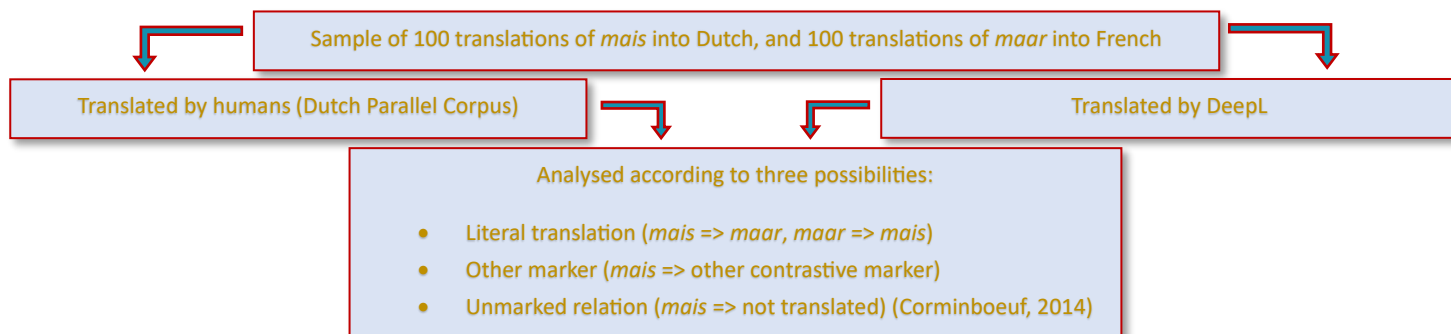
Nathanaël Stilmant (nathanael.stilmant@umons.ac.be). UMONS, FTI-EII, Service de Traductologie, Langue et Culture néerlandaises (TraLaNed)

1. Background: machine translation, a revolution for the translation profession

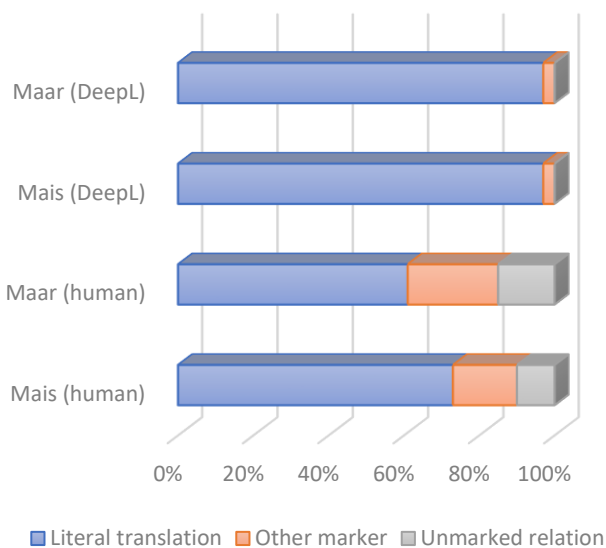


- ⇒ Are machine translations exactly the same as human translations?
- ⇒ Machine translation does not (yet) equal human creativity in certain respects (Vanroy, 2021), particularly when it comes to discourse markers (Meyer & Webber, 2013).

2. Methodology: comparison of the translations of the contrastive markers *mais* (FR) and *maar* (NL) from a human corpus and a machine corpus (DeepL)



3. Results: human translation strategies show greater variety (DeepL results collected in January 2023)



The few *maar* translated by other markers by DeepL appear in sentences where these *maar* are not alone, but are part of a sequence of markers (Luscher, 1993). This suggests that the choice of DeepL is influenced by the presence of such sequences, as in:

- i) Wie zou er niet geroerd zijn door een dergelijke spontane, **maar** toch 'doelgerichte grilligheid'?
- ii) Qui ne serait pas ému par une telle fantaisie spontanée **et pourtant** « volontaire » ?

On the other hand, some isolated *mais* have been translated by a marker other than *maar* by DeepL:

- a) En principe, l'épreuve aurait dû être disputée en deux manches. **Mais** en raison des sanctions imposées à l'époque à la Yougoslavie, la finale se limita à une seule rencontre qui eut pour cadre « Old Trafford ».
- b) In principe had de wedstrijd over twee legs gespeeld moeten worden. Door de sancties die Joegoslavië destijds opgelegd kreeg, bleef de finale **echter** beperkt tot één wedstrijd op Old Trafford.

4. Conclusions: discourse markers variety, a human specificity

- Human translations use literal translation far less often than DeepL ($\chi^2 = 56.9027$, $p = 0.00001$).
- DeepL never uses the strategy of the unmarked relation.
- In rare circumstances, DeepL shows variety and avoids literal translation. This seems to be facilitated by the presence of a sequence of discourse markers.

5. References

- Corminboeuf, G. (2014). *L'identification des relations de discours implicites : le cas de l'adversation*. Paper presented at 4e Congrès Mondial de Linguistique Française – CMLF 2014, Berlin.
- Luscher, J.-M. (1993). La marque de connexion complexe. *Cahier de linguistique française*, 14, 173-185.
- Meyer, T., & Webber, B. (2013). Implication of Discourse Connectives in (Machine) Translation. *Proceedings of the Workshop on Discourse in Machine Translation (DiscoMT)*, 19-26.
- Vanroy, B. (2021). *Syntactic Difficulties in Translation* (Doctoral dissertation, Ghent University).