What neural machine translations tell us about syntax in a contrastive linguistic perspective

In the realm of syntax, contrastive linguistics are not only concerned with detailed grammatical rules of different languages, but also with idiomatic preferences of sentence structure.

The present study investigates the degree to which neural machine translation (NMT) can imitate details of syntactic rules and produce correct and idiomatic syntactic output in different languages.

To date, machine translations are mostly evaluated according to general correctness and metagrammatical criteria (e.g. word choice, ellipsis or duplication of elements, pragmatics, text linguistics, punctuation, formatting; Popović 2018, Van Brussel et al. 2018, Krüger 2020).

The parallel corpus of EU legal acts was taken as the basis for comparison, as the texts are translated and revised by native speakers. A high degree of semantic convergence is required for these texts as they are equally authentic and legally binding in their 24 language versions.

For the study the English version of legal acts was translated into French and German by two neural machine translation systems. The results were compared to each other, to the source version and to the authentic French and German versions.

The findings were that English syntax was calqued in both French and German, that those calques lead to unidiomatic and sometimes ungrammatical constructions and that the machines could not cope with complex relators and information-value oriented syntax. However, the AI also produced sentences that were better structured than in the authentic version.

These results sharpen the awareness for details of syntactic rules and idiomaticity in different languages and contribute to the granularity of grammatical description, to translation studies including post-editing, and to the choice of input training material for NMT systems (cf. Klubička/Toral/Sánchez-Cartagena 2018, Isabelle/Cherry/Foster 2017), adapted to pairs of source and target languages.

Isabelle, Pierre/Cherry, Colin/Foster, George 2017: A Challenge Set Approach to Evaluating Machine Translation. In: Association for Computational Linguistics (ed.): *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing.* Copenhagen. 2486–2496.

Klubička, Filip/Toral, Antonio/Sánchez-Cartagena, Víctor M. 2018: Quantitative fine-grained human evaluation of machine translation systems: A case study on English and Croatian. *arXiv*:1802.01451v1.

Krüger, Ralph (2020): Propositional opaqueness as a potential problem for neural machine translation. In: Ahrens, Barbara/Beaton-Thome, Morven/Krein-Kühle, Monika (u.a.) (Hrsg.): *Interdependenzen und Innovationen in Translation und Fachkommunikation / Interdependence and Innovation in Translation, Interpreting and Specialised Communication.* Berlin: Frank & Timme. 261-278.

Popović, Maja 2018: Error classification and analysis for machine translation quality assessment. In: Moorkens, Joss/Castilho, Sheila/Gaspari, Federico/Doherty, Stephen (eds.): *Translation Quality Assessment. From Principles to Practice*. Cham, Switzerland: Springer.. 129-158

Van Brussel, Laura/Tezcan, Arda/Macken, Lieve 2018: A Fine-grained Error Analysis of NMT, PBMT and RBMT Output for English-to-Dutch. In: Proceedings of the *Eleventh International Conference on Language Resources and Evaluation.* European Language Resources Association (ELRA). 3799–3804. https://aclanthology.org/L18-1600/.