Translator Style and Machine Translation

Presenter(s): Mehmet Şahin ORCID 0000-0002-9077-1241, Tunga Güngör ORCID 0000-0001-9448-9422, **Ena Hodzik ORCID 0000-0001-5174-5596**, Sabri Gürses ORCID 0000-0002-6452-5426, Harun Dallı ORCID 0000-0001-8753-6024, Olgun Dursun ORCID 0000-0001-8753-6024, Zeynep Yirmibeşoğlu ORCID 0000-0001-5579-6489

This study, which is part of a scientific research project[1] funded by TüBİTAK (the Scientific and Technological Research Council of Turkey), takes an interdisciplinary approach investigating style in a corpus of Turkish literary translated texts for the purpose of 1) determining the distinctive measurable characteristics of translator style and 2) examining to what extent a customized MT system reflects a translator style when the system is trained on previous translations by the same translator. To answer the first research question, we applied Baker's (2000) corpus-based analysis of translator style and Leech and Short's (1981) method of literary style analysis and created a model of a particular Turkish translator's style. A classifier for translator attribution based on statistical language models and word embeddings will also be used to identify translator style. Our preliminary quantitative and qualitative corpus data analysis revealed distinctive stylistic features on lexical and discourse levels. Our research based on the biographical and historical data provided evidence for the impact of non-textual factors on the translator style. Taking these preliminary results into account, we will present the main stylistic features of our model and discuss expectations regarding the MT system training and testing stages of our study.

References:

Baker, M. (2000). Towards a methodology for investigating the style of a literary translator. *Target. International Journal of Translation Studies*, *12*(2), 241-266.

Leech, Geoffrey N., and Michael H. Short. (1981) 2007. *Style in Fiction: A Linguistic Introduction to English Fictional Prose.* London: Longman.

---

[1] The title of the project is "Literary Machine Translation to Produce Translations that Reflect Translators' Style and Generate Retranslations" and its grant number is 121K221.