

Abs-69

Lukasz Grabowski (Opole University)

Translation universals revisited: a corpus-driven comparison of translational and non-translational literary Polish

According to Baker (1993, 1995), corpus-based translation studies should focus, among others, on translation as a product by comparing comparable corpora of translational and non-translational texts. This article presents selected results of a corpus-driven comparison of the two custom-designed corpora of translational and non-translational Polish literary texts. The aim of the study was to find stylistic differences between the two corpora as well as identify traces, if any, of translation universals. Therefore, this study aims to examine T-universals (after Chesterman 2004) with emphasis on core patterns of lexical use, as proposed by Laviosa (1998), on the example of literary Polish.

The study was completed with the use of selected methods offered by corpus linguistics, namely descriptive statistics, distribution of n-letter words, frequency profiles, frequency spectra (after Baroni 2009) and computational stylometry (Classical Delta Distance, after Burrows 1986/2002). To that end, the corpora of translational and non-translational literary Polish were compiled. Each corpus includes approx. 500,000 word tokens and contains a selection of fiction novels published in Poland in the years 1937-2001 (non-translational corpus) and translated into Polish in the years 1950-2001 (translational corpus). Finally, with the use of custom-designed R script (Eder – Rybicki 2011), the texts were processed with the use of a multivariate methods [Principal Components Analysis (PCA) and Cluster Analysis (CA)] and presented graphically in order to illustrate relative distances between the translations and typical Polish novels in terms of 1000 the most frequent words (MFW) in each of the novels included in both corpora.

The results of the study showed that translational literary Polish does not conform to core patterns of lexical use found in typical literary Polish, in particular in terms of lexical variety among top-frequency words, word length and sentence length. Furthermore, PCA and CA showed that translations and typical literary Polish cluster together into two separate groups with considerable distance between them, which provides further evidence as to the existence of the levelling-out translation universal. The article concludes with suggestions and problems as regards research on translation universals with the use of language corpora.

Selected references

Baker, M. (1993). "Corpus linguistics and Translation Studies: Implications and applications". In M. Baker, G. Francis and E. Tognini-Bonelli (eds.), *Text and Technology. In Honor of John Sinclair*. Amsterdam: John Benjamins, 233-250.

Baker, M. (1995). "Corpora in translation studies: An overview and some suggestions for future research". *Target*, 7(2), 223-243.

Baroni, M. (2009). "Distributions in text". In A. Lüdeling and M. Kytö (eds.), *Corpus linguistics: An international handbook Volume 2..* Berlin and New York: Walter de Gruyter, 803-821.

Chesterman, A. (2004). "Beyond the particular". In A. Mauranen and P. Kuyamaki (eds.) *Translation Universals: Do they exist?*. Amsterdam: John Benjamins, 33-49.

Eder, M., Rybicki, J. (2011). "Stylometry with R". In: *Digital Humanities 2011: Conference Abstracts*, Stanford, CA (forthcoming).

Laviosa, S. (1998b). "Core patterns of lexical use in a comparable corpus of English narrative prose". In: *Meta*, 43(4), 557-570.

McEney, T., R. Xiao and Y. Tono (2006). *Corpus-based Language Studies: An Advanced Resource Book*. London and New York: Routledge.

Olohan, M. (2004). *Introducing Corpora in Translation Studies*. London and New York: Routledge.

Xiao, R. (2010). "How different is translated Chinese from native Chinese? A corpus-based study of translation universals". In: *International Journal of Corpus Linguistics* 15(1), 5-35.