

Can Machine Translation of Literary Texts Fool Stylometry?

Rybicki, Jan

jan.rybicki@uj.edu.pl
Uniwersytet Jagielloński, Poland

Introduction

Not so long ago, machine translation systems served as a butt of jokes among their human competitors. In fact, the “first and longer standing perspective” on MT dealing with literary texts tended “to see machine translation outputs as a kind of ‘found poetry’, ... often prized for their alienating, ungrammatical, anti-narrative or even comical effects” (Kenny and Winters 2021). But now MT has turned to Neural Networks/Deep Learning; and while its use in literary translation is still not widespread practice, this seems to be looming over the horizon.

Stylometric research on machine translation of literature is virtually non-existent. There is a growing body of qualitative case studies on MT in literary translation (Toral and Way 2015, 2018, Kenny and Winters 2020, Schmidt 2021, Zhang 2021), but only two texts mix quantitative/stylometric authorship attribution techniques with MT (Caliskan and Greenstadt 2012, Caliskan 2013).

Human literary translations have fared only slightly better in stylometric research. The notion of the invisible (literary) translator, criticized by Lawrence Venuti (1995), has been found to persist, to some extent, in stylometric studies (Rybicki 2012). Translators’ stylometric traces may be visible when they are followed in collaborative translations of the same text (Rybicki and Heydel 2013) or in texts of the same author translated by different humans (Rybicki 2006). Dependence of translators’ stylistic visibility was studied in relation to language combinations (Lee 2018) and individual translators’ styles (Burrows 2002a, Mastropierro 2018). Lee (2021) used a variety of linguistic features to show “a clear-cut divide in writing style between human and machine translators” and to observe that different MT systems are becoming more alike to each other in their output of literary translation.

The main focus of this experiment is to see whether machine translation systems now produce literary translations that do not depart in their stylometric features from those made by humans.

Methods

In this study, the originals and the English translations: human, Google Translate, and DeepL, are compared via stylometry based on very frequent words according to Burrows’s Delta (Burrows 2002b), combined with cluster analysis of distance measures (Smith and Aldridge 2011, Evert et al. 2017) and network analysis (Eder 2017), using the “stylo” package (Eder et al. 2016) for R (R Core Team 2021). This standard authorship attribution procedure will be adopted through a consensus cluster analysis and its network analysis visualizations.

Material

The 52 texts for this study were selected from classical 19th- and early 20th-century French literature and included 2-10 novels each by 10 writers. In the corresponding set of English translations, not a single author was translated by a single translator; also, no translator dealt with more than one original author. In all, 23 named translators were represented in the English set; 16 translations were anonymous. Full texts of the original novels were also translated into English with Google Translate and DeepL. Altogether, 157 human and machine translations were compared.

Results

Figure 1 shows a network analysis of cluster-analysed Cosine Delta distances between the French originals. As expected, the individual authors form their own communities.

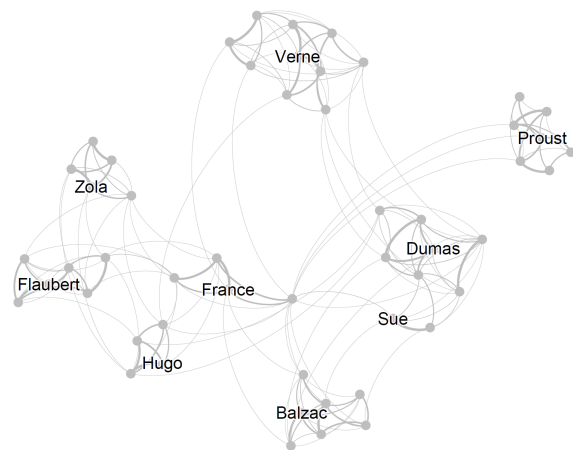


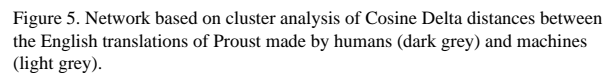
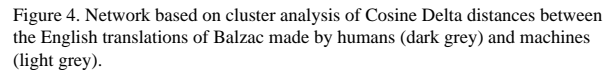
Figure 1. Network based on cluster analysis of Cosine Delta distances between the French originals.

The same diagram for the English translations shows the persistence of the signal of the original author (Figure 2); the networked patterns of similarity and difference between human and machine translations of the same texts come in two main forms.



A network graph visualization showing relationships between various authors and titles. The nodes are labeled with names like 'flaubert_bovary_gt', 'zola_nana_anon', etc., and are connected by curved edges representing relationships.

Sometimes, however, this pattern is broken. Translations of Balzac by Ellen Marriage form a sub-community of their own in the otherwise mixed group of other human and machine translations (Figure 4, bottom centre). And the human translations of Proust, especially those by C. K. Scott Moncrieff, depart even more clearly from their machine rivals (Figure 5).



Let's face it: machine translations of literature are slowly becoming a match for those made by humans, at least in terms of stylometric authorship (and/or translator) attribution. It cannot be a coincidence that only Proust and his version of the stream of consciousness characteristic in its highly complex sentences is too much for machine translations to successfully fool stylometry, while machine-translated traditional storytelling by Dumas is stylometrically much more similar to the human result, except perhaps for the absence of exoticization of the original *Monsieur* that human translators invariably used: "‘What!’ cried D’Artagnan, ‘your first witness is Mr. Porthos?’" In Proust, machine translation often loses the thread of the narrative, even in his most famous passage:

I carried a spoonful of tea on my lips where I had let a piece of madeleine soften. But at the very moment when the sip mixed with the crumbs of the cake touched my palate, I trembled, attentive to what was happening in me that was extraordinary.

2

Acknowledgement

The research has been supported by a grant from the Priority Research Areas

“DigiWorld” and “Heritage” under the Strategic Programme Excellence Initiative at Jagiellonian University.

Bibliography

- Burrows, John F.** (2002a): “The Englishing of Juvenal: computational stylistics and translated texts”, in: *Style* 36, 4: 677–99.
- Burrows, John F.** (2002b): “‘Delta’: a measure of stylistic difference and a guide to likely authorship”, in: *Literary and Linguistic Computing* 17, 3: 267–87.
- Caliskan, Aylin / Greenstadt, Rachel** (2012): “Translate Once, Translate Twice, Translate Thrice and Attribute: Identifying Authors and Machine Translation Tools in Translated Text”, in: IEEE (ed.): *ICSC '12 Proceedings of the 2012 IEEE Sixth International Conference on Semantic Computing, (Palermo, Italy)*, 121–25.
- Constantine, Peter** (2019): “Google Translate Gets Voltaire: Literary Translation and the Age of Artificial Intelligence”, in: *Contemporary French and Francophone Studies* 23, 4: 471–479, DOI: 10.1080/17409292.2019.1694798
- Eder, Maciej** (2017): “Visualization in stylometry: cluster analysis using networks”, in: *Digital Scholarship in the Humanities* 32, 1: 50–64.
- Eder, Maciej / Rybicki, Jan / Kestemont, Mike** (2016): “Stylometry with R: a package for computational text analysis”, in: *The R Journal* 8, 1: 107–121.
- Evert, Stefan / Proisl, Thomas / Jannidis, Fotis / Reger, Isabella / Pielström, Steffen / Vitt, Thorsten** (2017): “Understanding and explaining Delta measures for authorship attribution”, in: *Digital Scholarship in the Humanities* 32 suppl. 2: ii4–ii16.
- Kenny, Dorothy / Winters, Marion** (2021): “Machine Translation in Literary Translation: A Case Study”, in: *Computer-Assisted Literary Translation Conference CALT2021@Swansea (11-13 May 2021) Book of Abstracts*, Swansea: Swansea University.
- Lee, Changsoo** (2018): “Do language combinations affect translators’ stylistic visibility in translated texts?”, in: *Digital Scholarship in the Humanities* 33, 3: 592–603.
- Lee, Changsoo** (2021): “How do machine translators measure up to human literary translators in stylometric tests?”, in: *Digital Scholarship in the Humanities* 37, 3: 813–829.
- Mastropierro, Lorenzo** (2018): *Corpus Stylistics in Heart of Darkness and its Italian Translations*. London: Bloomsbury.
- R Core Team** (2021): *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rybicki, Jan** (2006): “Burrowing into Translation: Character Idiolects in Henryk Sienkiewicz’s Trilogy and its Two English Translations”, in: *Literary and Linguistic Computing* 21: 91–103.
- Rybicki, Jan** (2012): “The Great Mystery of the (Almost) Invisible Translator. Stylometry in Translation”, in: Oakes, Michael P. / Ji, Meng (eds): *Quantitative Methods in Corpus-Based Translation Studies*, Amsterdam: John Benjamins 231–48.
- Rybicki, Jan** (2021): “Stylometry 0, Machine Translation 1. Deep Learning Based MT Scores Important Away Win”, *Computer-Assisted Literary Translation Conference CALT2021@Swansea (11-13 May 2021) Book of Abstracts*, Swansea: Swansea University.
- Rybicki, Jan / Heydel, Magda** (2013): “The Stylistics and Stylometry of Collaborative Translation: Woolf’s Night and Day in Polish”, in: *Literary and Linguistic Computing* 28, 4: 708–17.
- Schmidt, Bradley** (2021): “Ghosts in the Machine”, *Computer-Assisted Literary Translation Conference CALT2021@Swansea (11-13 May 2021) Book of Abstracts*, Swansea: Swansea University.
- Smith, Peter W. H. / Aldridge, William** (2011): “Improving Authorship Attribution: Optimizing Burrows’ Delta Method”, in: *Journal of Quantitative Linguistics* 18, 1: 63–88.
- Toral, Antonio / Way, Andy** (2015): “Machine-assisted translation of literary text: a case study”, in: *Translation Spaces* 4, 2: 240–267.
- Toral, Antonio / Way, Andy** (2018): “What level of quality can Neural Machine Translation attain on literary text?”, in: Moorkens, Joss / Castilho, Sheila / Gaspari, Federico / Doherty, Stephen (eds): *Translation Quality Assessment: From Principles to Practice*. Heidelberg: Springer International Publishing 263–287.
- Venuti, Lawrence** (1995): *The Translator’s Invisibility: A History of Translation*. Oxford: Routledge.
- Zhang, Shuyin** (2021): “On discussion of the machine translation of Chinese Xianxia novels”, *Computer-Assisted Literary Translation Conference CALT2021@Swansea (11-13 May 2021) Book of Abstracts*, Swansea: Swansea University.