

Tracing the invisible translator: stylistic differences in the Dutch translations of the oeuvre of Swedish author Henning Mankell

Wijers, Martje

m.wijers@uva.nl

University of Amsterdam, Netherlands, The

Baker (2000) stated that very little attention has been paid to personal style of translators in translation studies, due to the persistent idea that style is a characteristic of the original text only and that the style of a good translator is invisible. Since then, some attempts have been made to investigate translator's style from a stylometric perspective. Rybicki (2012) applied stylometric authorship attribution methods, based on the Delta procedure for authorship attribution (Burrows, 2002) to compare translations in multiple corpora containing Polish, English and French translations. The Delta method is based on the most frequent words (MFWs) in the entire corpus. This method has proven to be highly successful in authorship attribution. However, Rybicki (2012) found that the Delta method could not identify different translator clusters, as it rather clusters texts by author. From these results, it could be concluded that, based on Burrow's delta, translators are indeed invisible when it comes to style.

However, when translations of texts by the same author are compared, translators can in some cases be detected (Rybicki 2012, Rybicki & Heydel 2013). Furthermore, Rybicki (2012) proposes other stylometric methods, like Burrow's Zeta and Iota (Burrows, 2006) to investigate translator's style, because they focus on middle and low frequency words and might be better suited for this purpose. Although recent studies have tried to solve the problem of translator attribution, they either focus on n-grams using machine learning (Lynch & Vogel 2018; Ríos-Toledo et al., 2022) or on MFWs (Fang & Liu, 2022). There are no studies, to my knowledge, that have tested Rybicki's claim that Burrow's Zeta and Iota are useful for looking at the translator's style. Burrow's Zeta (Burrows, 2006; Hoover, 2022) can be used to find distinctive words or n grams between texts or groups of texts. This way, translations from one translator can be compared to translations from other translations in the same genre and by the same author.

In this computational study, the oeuvre of one single Swedish author Henning Mankell is scrutinized using stylometric methods, mainly Burrow's zeta, and his original works are then compared to their Dutch translations to investigate to what extent the results deviate from the same analyses in translated texts and to find out to what extent the style is influenced by the translator. The work of this one specific author is chosen, because it is hypothesized that this makes it easier to detect the style of the translator. Furthermore, the language combination Swedish-Dutch has not been compared in a similar stylometric study before and because the languages are closely related it is hypothesized that the translator is particularly difficult to find compared to more distinct language combinations (Lee, 2018).

The corpus compiled for the current study consists of 32 Swedish books written by Mankell in four genres: crime-fiction (N=15), literary novels (N=11), children's books (N=4) and non-fiction (N=2). For comparison purposes, ten books by other best-selling Swedish writers were added to the corpus. The translation corpus contains 42 translations by 14 translators of all the above-mentioned works by Henning Mankell and other Swedish writers into Dutch.

A stylometric analysis was performed on the Mankell corpus using the R package Stylo (Eder et al 2016). The Stylo package automatically compiles a list of MFWs in the entire corpus and can compare word frequency patterns in the various texts, based on various procedures for authorship attribution. The results of the cluster analysis and the principal components analysis of the Swedish corpus show that novels cluster by author and genre rather than by translator, as expected by earlier findings. However, the results from the same stylometric analyses on the translated Dutch corpus show that the distances between Mankell and other authors are much smaller compared to the Swedish corpus. This implies that some of the author-specific style traits get lost in translation. A comparative analysis of texts by a specific translator to texts by other translators using Burrow's Zeta, returns remarkably few function words, but is able to pinpoint some individual, lexical translator traits.

Other measures, such as lexical density (Standardized Type Token Ratio) and mean sentence length (as proposed by Baker, 2000) are also applied, as well as a Zeta analysis on n-grams. The results show that the translations score systematically lower than the Swedish originals on lexical density, whereas sentences are generally longer in the Dutch translations. By combining the various measures, the translator's style can be unveiled.

Bibliography

Burrows, John (2002): "'Delta': a measure of stylistic difference and a guide to likely authorship." *Literary and linguistic computing* 17.3: 267-287.

Burrows, John (2006): All the way through: testing for authorship in different frequency strata. *Literary and Linguistic Computing* , 22(1), 27-47.

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

Eder, Maciej / Rybicki, Jan / Kestemont, Mike (2016): Stylo-metry with R: a package for computational text analysis. *R Journal* , 8(1): 107-21.

Fang, Yu / Liu, Haitao (2022): Seeing various adventures through a mirror: Detecting translator's stylistic visibility in Chinese translations of Alice's Adventure in Wonderland. *Digital Scholarship in the Humanities* . Volume 38, Issue 1.

Lee, Changsoo (2018): Do language combinations affect translators' stylistic visibility in translated texts?. *Digital Scholarship in the Humanities* , 33 (3), 592-603.

Lynch, Gerard / Vogel, Carl (2018): The translator's visibility: Detecting translatorial fingerprints in contemporaneous parallel translations. *Computer Speech & Language* , 52 , 79-104.

Mohamed, Emad / Sarwar, Raheem / Mostafa, Sayed (2022): Translator attribution for Arabic using machine learning. *Digital Scholarship in the Humanities*. 13 October 2022.

Ríos-Toledo, Germán / Posadas-Durán, Juan Pablo Francisco / Sidorov, Grigori & Castro-Sánchez, Noé Alejandro (2022): Detection of changes in literary writing style using N-

grams as style markers and supervised machine learning. *Plos one*, 17 (7).

Rybicki, Jan (2012): The great mystery of the (almost) invisible translator. *Quantitative Methods in Corpus-Based Translation Studies: A practical guide to descriptive translation research*, 231.

Rybicki, Jan / Heydel, Magdalena (2013): The stylistics and stylometry of collaborative translation: Woolf's Night and Day in Polish. *Literary and Linguistic Computing*, 28 (4), 708-717.