

Original, Translation, Inflation. Are All Translations Longer than Their Originals?

Rybicki, Jan

jkrybicki@gmail.com

Pedagogical University, Krakow, Poland

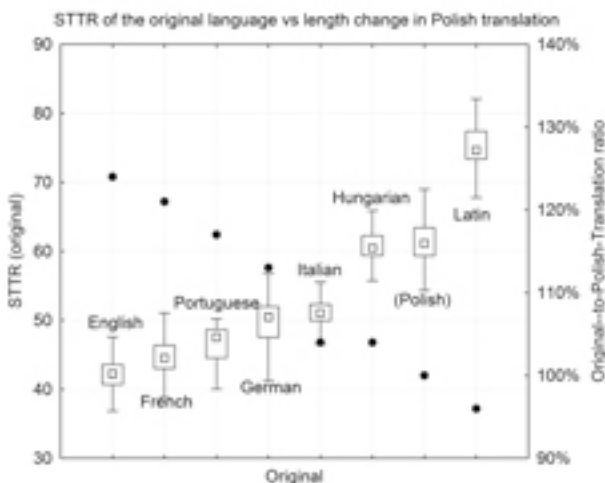
It is a truth almost universally acknowledged, at least among translator service providers, that some languages take fewer words to express the same thing than some other languages, to the extent that translator remuneration is often calculated accordingly. To further paraphrase Jane Austen and John Burrows: this truth is so well fixed in the minds of the general translating community that the scant reports pointing to the contrary – or, at least, to a possibility that this effect might be exactly contrary to expectations – are either ignored or appear in the wrong journals (Rybicki, 2006). While this problem is not entirely ignored by traditional translation studies, it is usually dealt with as an aside in publications where this discipline meets corpus linguistics to define and study translator style (Baker, 1993, 1996, 2000), or applied to no more than two languages, very few texts and, more often than not, small sample sizes (Englund Dimitrova, 1994, 2003, Pápai, 2004), or oriented to point out differences between two translations rather than original and translation (Rybicki, 2009). Theoretical considerations are just as unsatisfactory. Differences in the level of inflection of the two languages are usually seen as the reason for the differences in length between the native and the foreign version of the same text; the rare exception, i.e. a more or less positive statement on the subject, has been made by George Steiner: “translations are inflationary” (Steiner, 1978) in a discussion of explicitation, one of the so-called translation universals (Baker, 1993, 1996). Still, while explicitation is a mechanism that certainly does involve using more text in the target language to denote less text in the original, it is not clear whether Steiner had any specific textual unit in mind that would undergo inflation – as opposed to another possibility, the inflation of meaning.

Indeed, it is even less clear whether mere difference in the number of words – the first and reflexive approach most stylometrists would take – between a novel in one language and another novel, the former's translation, is at all of any scholarly interest; it is quite possible that what matters more is the increase (or decrease) in the number and/or the length of, say, sentences. Even then, however, the differences could be a simple consequence of the divergent linguistic systems and the whole problem should be left at that.

It is almost a tradition that, faced with such theoretical quandaries, members of our community turn to empirical practicalities, to experiment – and this is exactly what this paper does. Using a series of fairly extensive bilingual corpora or, simply speaking, combinations of original and translation (and, in some cases, another, and yet another translation of the same text) in a variety of source and target languages, the study compares the sizes, establishes their patterns and their statistical significance (with *z*-scores). The corpora in question include: English translations of Polish novels by Henryk Sienkiewicz; Polish translations of American, English, French, German and Italian prose (including the interesting sub-corpora of translations of Tolkien and of translations by the author of this paper); French and Polish translations of Shakespeare; Polish and English translation of Latin prose, Portuguese translations of English prose.

The results do not paint a uniform picture. While expected general trends can be observed in size variation between pairs of languages, the discrepancies in “inflation rate” between certain rival translations into the same language at times hide any stable “language-to-language” effect. This effect has been hypothetically ascribed at first to differences between inflected (agglutinative) and analytic languages. While this would be difficult to prove, at the same time – barring such extreme cases of translator logorhea as W.S. Kuniczak's famously overflowing translation of Henryk Sienkiewicz's historical romances, where the translation-to-original ratio reaches the vertiginous heights of 170%, the record value in the entire project – some correlation has been observed not so much between the general degree of inflection of a given language as between standardized type-

token ratios in each of the studied individual-language corpora. Thus, although it would be too much to say that STTR is a good measure of a language's inflection, the general trend in STTR ranges observed in each of the corpora used in this study corresponds fairly well to the *inverted* order of languages exhibiting difference between original and translation (see Figure below): translations into English tend to be longer than their Polish originals; Polish translations are shorter than original English novels; most translations of Latin prose tend to be longer than the originals, and so forth. With an important caveat: it only takes an overambitious, overzealous or pathologically lazy translator, or an unscrupulous publisher, to alter this pleasant image beyond recognition.



Standardized Type-Token Ratio (Box & Whisker) and Original-to-Translation Ratio (Scatterplot) in Selected Prose Corpora

References

- Baker, M.** (1993). 'Corpus linguistics and translation studies: Implications and applications'. *Text and Technology: In honour of John Sinclair*. Baker, M., Francis, G., Tognini-Bonelli, E. (eds.). Amsterdam: John Benjamins, pp. 17-45.
- Baker, M.** (1996). 'Corpus-based translation studies: The challenges that lie ahead'. *Terminology, LSP and Translation: Studies in language engineering, in honour of Juan C. Sager*. Somers, H. (ed.). Amsterdam: John Benjamins, pp. 175-186.

Baker, M. (2000). 'Towards a methodology for investigating the style of a literary translator'. *Target*. **12**: 241-266.

Pápai, V. (2004). 'Explicitation – A universal of translated text?'. *Translation Universals. Do they exist?*. Mauranen, Kujamäki (eds.). Amsterdam – Philadelphia: John Benjamins, pp. 143-164.

Englund Dimitrova B. (1994). *Statistical Analysis of Translations (On the basis of translations from and to Bulgarian, Russian and Swedish)*. Scandinavian Working Papers on Bilingualism. V. 9, pp. 87-103.

Englund Dimitrova B. (2003). 'Explicitation in Russian-Swedish translation: sociolinguistic and pragmatic aspects'. *Swedish Contributions to the Thirteenth International Congress of Slavists, Ljubljana, 15-21 August 2003*. Englund Dimitrova B., Pereswetoff-Morath, A. (eds.). Lund: Lund University, pp. 21-31.

Rybicki, J. (2006). 'Burrowing into Translation: Character Idiolects in Henryk Sienkiewicz's Trilogy and its Two English Translations'. *Literary and Linguistic Computing*. **21(1)**: 91-103.

Rybicki, J. (2009). 'Liczenie krasnoludków. Trochę inaczej o polskich przekładach trylogii Tolkiena'. *Po co ludziom krasnoludki?*. Warszawa, 2009.

Steiner, G. (1978). *After Babel. Aspects of Language and Translation*. Oxford: Oxford University Press V. 253, reprinted, 1992.