

Investigating Lexical Patterning in Translated Brazilian Portuguese: A Corpus-based Study

Carmen Dayrell

carmen_dayrell(a)hotmail.com

ABSTRACT

The primary objective of this study is to develop a suitable corpus-based research methodology for comparing lexical patterning in translated and non-translated language. The main issue under investigation is whether collocational patterns tend to be less diverse in translated texts in comparison with non-translated texts of the same language. Diversity is analysed in quantitative as well as qualitative terms.

The analysis is based on a comparable corpus of Brazilian Portuguese (BPC), which has been compiled as part of this research project. The corpus consists of two separate subcorpora designed according to the same criteria and specifications, one made up of translated Brazilian Portuguese and the other consisting of non-translated Brazilian Portuguese. BPC comprises two different text genres – fiction and self-help – and hence it is composed of four subcorpora: translated fiction, non-translated fiction, translated self-help and non-translated self-help. Collocational patterns of translated and non-translated texts are compared within each genre; that is to say, translated fiction is compared with non-translated fiction, and translated self-help with non-translated self-help. BPC contains approximately 2 million words in total, consisting of 1 million words of translated Brazilian Portuguese and 1 million words of non-translated Brazilian Portuguese. Each subcorpus contains approximately 0.5 million words.

The results indicate that collocational patterns tend to be less diverse in translated texts in comparison with non-translated texts of the same language, irrespective of the text genre. It is also found that differences between translated and non-translated texts are more pronounced in the fiction than in the self-help genre. The latter seems to reflect the peculiarity of the genre within the context of Brazilian Portuguese; self-help is in itself a translated genre.

KEYWORDS: collocations, corpus linguistics, lexical patterning, Portuguese.

Completion of thesis

Place: University of Manchester

Year: 2005

Supervisor: Mona Baker