Towards a Corpus-Based Methodology for Investigating Word Meaning in Translated Texts

Carmen Dayrell[1]

Abstract

Based on claims in the literature that relevant differences can be found in the lexical choices of translated and non-translated language, this paper sets out to examine potential differences between translated and non-translated texts of the same language with respect to the way words are combined in order to convey meaning. The primary issue under investigation is whether these two varieties of the same language display different tendencies and preferences in relation to the different senses of a given word-form. Here, meaning is analysed on the basis of the surrounding cotext of the word-form in question, following the suggestion that meaning correlates with particular grammatical and lexical choices. A small-scale study is carried out for the Portuguese lexical item 'passar' with the view to illustrating the results that may be obtained by such comparison. Data is drawn from a comparable corpus of Brazilian Portuguese which consists of two separate subcorpora: one made up of translated Brazilian Portuguese and the other consisting of texts which have been originally written in Brazilian Portuguese. Both subcorpora are designed according to the same criteria and specifications. The corpus 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. The comparison is made between translated and non-translated texts within the same genre. The results indicate that translated and non-translated texts seem to show different preferences towards the use of the different senses of a given word-form, irrespective of the text genre. In addition, for some senses, it is also found that meaning is signalled differently across translated and non-translated texts of the same genre.

---

[1] NILC (Research Centre of Computational Linguistics), University of São Paulo, Brazil
   e-mail: c_dayrell@yahoo.com