

A corpus-based multivariate analysis of linguistic norm-adherence in audiovisual and written translation

Lynn Prieels and Gert De Sutter

Research unit EQTIS (Empirical and Quantitative Translation and Interpreting Studies)
Department of Translation, Interpreting and Communication – Ghent University
Ghent, Belgium

Previous research within corpus-based translation studies has shown that written translations are more normalized than their source texts or comparable non-translated texts (Baker 1993). However, recent studies have repeatedly demonstrated that this standardization tendency depends on contextual parameters such as register, source language and target audience (e.g. Delaere & De Sutter 2013; Kruger & van Rooy 2012). In our study this vexed question is sent in a new, largely unexplored direction, viz. audiovisual translation (AVT). Although AVT is a widely investigated discipline within Translation Studies, research that focuses on linguistic variability in audiovisual translation is relatively scarce. Most of the attention went to the exploration of the general strategies that are used to cope with the information load in the original text (e.g. Barambones Zubiria 2012) and to specific linguistic features in AVT (e.g. Baños 2013).

The present study measures linguistic norm-adherence in Belgian Dutch written and audiovisual translation. More particularly, it is investigated (i) to what extent Flemish subtitlers prefer non-standard Belgian Dutch variants rather than General Standard Dutch variants and (ii) to what extent their choices differ from those made by translators of written texts and by authors of original, non-translated texts. Furthermore, we explain the subtitlers' linguistic behavior through the parameters program genre and source language. In order to achieve that goal, we gathered sets of (lexical and grammatical) norm-related linguistic variables and extracted them from two corpora: (i) the subtitle data were extracted from the SoNaR-corpus, a 500-million word balanced reference corpus for contemporary (1954-present) written Dutch (Schuurman et al. 2010) and (ii) the written text material was extracted from the Dutch Parallel Corpus (Macken et al. 2011), a bidirectional parallel corpus with (Belgian and Netherlandic) Dutch as a source language and as a target language. Using profile-based correspondence analysis (Plevoets, 2008), we measured linguistic distances between the parameters and their interactions and visualized them in a two-dimensional plot.

The results reveal significant differences between subtitles and written translations, and between subtitles and original texts. More specifically, it is shown that subtitles hold a middle position between written translations and non-translations, as the subtitle data contained significantly more non-standard Belgian Dutch variants than regular written translations but less than original Dutch texts. In-depth analyses pointed out that linguistic choices in subtitles are determined by both the source language and by the program genre. On the one hand, it is shown that the intralingual subtitles of Flemish speakers contain more non-standard Belgian Dutch than the interlingual subtitles of English speakers and the intralingual subtitles of Netherlandic Dutch speakers. On the other hand, certain television genres (e.g. fiction and comedy) tend to encourage the use of non-standard Belgian Dutch in the subtitles, whereas in other genres (documentaries and children's television) the subtitles mainly contain standard language. Based on these results, we can conclude that Flemish subtitles tend to be normalized, but in a less extreme way than regular written translations, due to the fact that they are (heavily edited) translations on the one hand (stimulating norm-adherent behavior), and written reproductions of (spontaneous) spoken language with its colloquial features on the other hand (stimulating non-standardizing behavior).

References

- Baker, M. 1993. "Corpus Linguistics and Translation Studies. Implications and Applications." In *Text and Technology. In Honour of John Sinclair*, ed. by Mona Baker, Gill Francis and Elena Tognini-Bonelli, 233-250. Amsterdam: John Benjamins.
- Baños, R. 2013. "'That is so cool': investigating the translation of adverbial intensifiers in English-Spanish dubbing through a parallel corpus of sitcoms." In *Perspectives: Studies in Translatology*. 21, no. 4: 526-542.
- Barambones Zubiria, J. 2013. *Mapping the Dubbing Scene. Audiovisual Translation in Basque Television*. Oxford: Peter Lang.
- Delaere, I., & De Sutter, G. 2013. "Applying a Multidimensional, Register-sensitive Approach to Visualize Normalization in Translated and Non-translated Dutch. Belgian Journal of Linguistics." In: M.-A. Lefer & S. Voegleer (eds.) *Interference and normalisation in genre-controlled multilingual corpora*, Vol. 27. Philadelphia/Amsterdam: Benjamins.
- Kruger, H. & B. van Rooy 2012. "Register and the features of translated language." In *Across Languages and Cultures*. 13, no 1: 33–65.
- Macken, L., De Clercq, O. & H. Paulussen 2011. Dutch Parallel Corpus: A balanced, copyright-cleared parallel corpus. *Meta: Translators' Journal*. 56, no. 2: 374-390.
- Plevoets, K. 2008. *Tussen spreek-en standaardtaal: een corpusgebaseerd onderzoek naar de situationele, regionale en sociale verspreiding van enkele morfo-syntactische verschijnselen uit het gesproken Belgisch-Nederlands*. Doctoral dissertation, Katholieke Universiteit Leuven. Faculteit Letteren.
- Schuurman, I., Hoste, V. and Monachesi, P., 2010, May. Interacting Semantic Layers of Annotation in SoNaR, a Reference Corpus of Contemporary Written Dutch. In *LREC*.