

Abs-41

Isabelle Delaere, Koen Plevoets and Gert De Sutter (University College Ghent, Ghent University)

COMURE: corpus-based, multivariate research investigating register variation between translated and non-translated Belgian Dutch.

Since the 1990s, research in the field of translation studies has not only been focusing on the relation between source and target texts, but also on the relation between translated language versus non-translated language (Baker 1993, 2004). Corpus-based research in translation studies has shown that translated texts differ in a surprisingly systematic manner from non-translated texts resulting in, for example, lexical (Laviosa 1998, Kemppanen 2004, Tirkkonen-Condit 2004) and grammatical differences (Puurtinen 2003, Olohan & Baker 2000). It is, however, remarkable that little research has been carried out to examine the differences between registers in translated versus non-translated language.

We want to investigate (i) to what extent the linguistic choices made in translations are register dependent and (ii) whether the linguistic differences between registers in translated texts are identical to those between these same registers in non-translated texts. To this end, we use the Dutch Parallel Corpus (DPC), a 10-million-word, high-quality, sentence-aligned parallel corpus for the language pairs Dutch-English and Dutch-French, with Dutch as the central language. It consists of the following registers: fiction, non-fiction, journalistic texts, instructive texts, administrative texts and external communication. (Macken et al. 2011).

We selected a set of linguistic profiles, viz. sets of alternatives with the same meaning and function such as *jij kunt* vs. *jij kan* (you can), *u heeft* vs. *u hebt* (you have) en *heeft gedaan* vs. *gedaan heeft* (have done). On the basis of the frequency of these alternatives in the DPC, we used profile-based correspondence analysis (Plevoets 2008) in order to measure and plot the linguistic distances between the various registers. The general idea behind this technique is that the linguistic distances between registers and translated vs. non-translated texts grow as a function of the extent to which linguistic choices differ.

Preliminary results show that there is a significant distance between translated and non-translated Dutch. More specifically, we observe a tendency towards the use of rather formal language for translated texts and a tendency towards the use of rather informal language for non-translated texts. Furthermore, we see that the linguistic choices made by authors or translators are register dependent, and, for external communication and journalistic texts, these choices are significantly different for translated versus non-translated texts.

Baker, M. (1993). Corpus linguistics and translation studies. Implications and applications. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and technology. In honour of John Sinclair*. (pp. 233-250). Philadelphia/Amsterdam: John Benjamins.

Baker, M. (2004). A corpus-based view of similarity and difference in translation. *International Journal of Corpus Linguistics*, 9(2), 167-193.

Kemppanen, H. (2004). Keywords and ideology in translated history texts: A corpus-based analysis. *Across Languages and Cultures*, 5(1), 89-107.

Laviosa, S. (1998). Core patterns of lexical use in a comparable corpus of English narrative prose. *Meta*, 43, 557-570.

Macken, L., De Clercq, O., & Paulussen, H. (2011). Dutch parallel corpus: A balanced copyright-cleared parallel corpus. *Meta*, 56(2).

Olohan, M., & Baker, M. (2000). Reporting that in translated English: Evidence for subconscious processes of explicitation? *Across Languages and Cultures*, 1(2), 141-158.

Plevoets, K. (2008). *Tussen spreek- en standaardtaal. Een corpusgebaseerd onderzoek naar de situationele, regionale en sociale verspreiding van enkele morfosyntactische verschijnselen uit het gesproken Belgisch-Nederlands*. Doctoral dissertation, Katholieke Universiteit Leuven, Leuven.

Puurtinen, T. (2003). Genre-specific features of translationese? Linguistic differences between translated and non-translated Finnish children's literature. *Literary and Linguistic Computing*, 18(4), 389-406.

Tirkkonen-Condit, S. (2004). Keywords and ideology in translated history texts: A corpus-based analysis. In A. K. Mauranen, Pekka (Ed.), *Translation universals. Do they exist?* (pp. 177-184). Amsterdam/Philadelphia: John Benjamins.