

Translation as specialised language use. A probabilistic view of linguistic characteristics of translations

Stella Neumann, RWTH Aachen University

Translation is simply language use, albeit specialised language use. If it wasn't, translations would be rejected as incomprehensible in the target language (TL). In many cases, the translation will be covert (House 1997) intended to work in the same way as a non-translated text. Against this background, we would expect translations to be indistinguishable from texts produced in the TL without an anterior text (Halverson 2013) in the source language. Moreover, the (non-translated) texts produced in the TL will also include other types of texts produced under conditions of language contact, i.e. involving a second language in addition to the TL. It therefore appears implausible to expect any linguistic peculiarities in (covert) translations. Yet, many computational studies have reported high accuracies in classifying translated texts based on their linguistic features (for an early example, see Baroni and Bernardini 2006). So, there must be something that makes translations easy to spot for the computer. In this paper, I will review the corpus-based approach to translation adopting a probabilistic view of language. Translation is a specialised form of language use because of its link to an anterior text: information linguistically encoded in the anterior text has to be re-coded in a different language system (Hansen-Schirra and Steiner 2012; similarly Halverson 2013). Hansen-Schirra and Steiner claim that translation is arguably the only type of text production which is linked to such a previous encoding – at least when excluding “[a]ny weaker form of multilingual text production in the sense of producing target context-adapted re-creation” (2012, 261). If their claim is right, precisely this link may be the reason for the specific distribution of linguistic features to which the computational classification task responds. Pressure on the translator to re-encode the anterior text's meaning and wording might lead to observable linguistic differences from non-translated TL texts, while attempting to cover up the fact that someone else has expressed this meaning previously in a different language might result in differences from the encoding of the anterior text and, more generally, the source language.

The by now widely accepted view of language as a dynamic, open, that is, probabilistic system (e.g. Halliday 1991, Beckner et al. 2009, see also Toury 2004) offers an explanation for the uneven, yet systematic distribution of linguistic features across texts produced under different conditions. This view implies a paradigmatic conceptualisation of language according to which language users have a range of more or less likely linguistic options at their disposal for expressing a certain meaning. Since translation is language use it is subject to the same types of systematic variation in probabilities that also apply to non-translated texts. The translator as language user consequently also chooses between various options even if the default translation is blocked, resulting in translation shifts, thus explaining variation between translations. The choice is systematically influenced by factors such as situational context, but also social and cognitive factors. More specifically, the findings of the above mentioned computational studies suggest that the influence of various factors applies differently in translations as compared to non-translated texts. Consequently, as suggested by De Sutter and Lefer (2020), the task of empirical translation studies is to ascertain the influence of the multiple factors that condition the outcome of the translation process. From a probabilistic point of view, the complex influence of factors affecting translational language use means that the effect is discernible, but weak. Crucially, this also means that translations will be gradually, not categorically different from other types of text. This has consequences not only for the corpus methodology in translation studies, thus aligning with De Sutter and Lefer's (2020) call for methodological innovation, but also for the way we interpret the results of corpus studies.

References

- Baroni, Marco, and Silvia Bernardini. 2006. 'A New Approach to the Study of Translationese: Machine-Learning the Difference between Original and Translated Text'. *Literary and Linguistic Computing* 21 (3): 259–74.
- Beckner, Clay, Richard Blythe, Joan Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman, and Tom Schoenemann. 2009. 'Language Is a Complex Adaptive System: Position Paper'. *Language Learning* 59 (s1): 1–26.
- De Sutter, Gert, and Marie-Aude Lefer. 2020. 'On the Need for a New Research Agenda for Corpus-Based Translation Studies: A Multi-Methodological, Multifactorial and Interdisciplinary Approach'. *Perspectives* 28 (1): 1–23.
- Halliday, M. A. K. 1991. 'Towards Probabilistic Interpretations'. In *Functional and Systemic Linguistics. Approaches and Uses*, edited by Eija Ventola, 39–61. Berlin, New York: Mouton de Gruyter.
- Halverson, Sandra L. 2013. 'Implications of Cognitive Linguistics for Translation Studies'. In *Cognitive Linguistics and Translation Advances in Some Theoretical Models and Applications*, edited by Ana Rojo and Iraide Ibarretxe-Antunano, 33–74. Berlin, Boston: de Gruyter.
- Hansen-Schirra, Silvia, and Erich Steiner. 2012. 'Towards a Typology of Translation Properties'. In *Cross-Linguistic Corpora for the Study of Translations*, Silvia Hansen-Schirra, Stella Neumann, and Erich Steiner, 255–79. Berlin: De Gruyter.
- House, Juliane. 1997. *Translation Quality Assessment. A Model Revisited*. Tübingen: Gunter Narr Verlag.
- Toury, Gideon. 2004. 'Probabilistic Explanations in Translation Studies. Welcome as They Are, Would They Qualify as Universals?' In *Translation Universals. Do They Exist?*, edited by Anna Mauranen and Pekka Kujamäki, 15–32. Amsterdam/Philadelphia: John Benjamins Publishing Company.