

Fabio Alves, Adriana Pagano¹, Silvia Hansen-Schirra², Stella Neumann, Erich Steiner³

This paper reports on a pilot study bringing together researchers with previous experience in process-oriented translation research (cf. Alves & Gonçalves 2003, Hansen 2003, Pagano et al. 2005) and in product-oriented corpus-based research (cf. Hansen-Schirra et al. 2007, Hansen-Schirra et al. 2006, Steiner 2008). The aim of the research is to arrive at a more comprehensive model of translation-oriented language processing, and at explanations of the different types of data. Empirically-based translation research has so far been developed within two major self-standing approaches: corpus-based work on properties of translated texts (*product*) and experimental studies of translators' expert performance (*process*). Recent advances in corpus architecture and multi-level corpus querying and an increasing incorporation of methods from psycholinguistics and cognitive science into process-oriented research (cf. Carl et al. 2008) point to a desired combination of the ecological validity of corpus studies with a more direct insight into processing efforts in psycholinguistic experiments.

The study combines the multi-layer annotation and alignment used for an English-German corpus of originals and translations in both languages with key-logging and eye-tracking of the translation process by two expert subjects translating from German into English. The source text used in the pilot study is an excerpt from a corporate webpage, a register that is included in the annotated corpus.

We investigate two types of phenomena - *units of translation* and *grammatical metaphor* - to explain properties of translations that can be tracked in the process. The research architecture is structured as follows:

- *Units of translation*: The annotation and alignment of the corpus allows identifying segments without corresponding translations ("empty links") and aligned segments which are not part of the same aligned segments on a higher level, e.g. aligned words in different grammatical functions ("crossing lines"). The process-oriented research analyses pauses in key-logging to identify translation units as instances of foci of attention. The aligned segments are then mapped onto the cognitive units. Eye-tracking is used to identify patterns of eye fixations that can be correlated to text processing generating empty links and crossing lines.
- *Grammatical metaphor* (i.e. semantic units mapped onto different grammatical units): Grammatical shifts are detected by querying aligned segments in the corpus (cf. Čulo et al. 2008). Drawing on key-logging and eye-tracking data, foci of attention (and potential processing problems) can be identified in instances of grammatical metaphor. We thus obtain insights into the routes in unpacking and repacking meanings. Eye-tracking helps identifying the elements in the source text that trigger this process.

The initial findings show that chunks of processing as divided by pauses, i.e. cognitive units, tend to be mapped onto phrases in grammatical terms. The macro unit, i.e. the unit between pauses that incorporates a sequence of steps taken by the translator to the final output (Alves & Vale, forthcoming), provides the frame for tracking empty links, crossing lines and (de-)metaphorization. On the basis of these initial findings we will set up a larger research design that will provide results on patterns of processing in combination with patterns of text production.

References

Alves, F. (ed.) 2003. *Triangulating translation: perspectives in process oriented research*. Amsterdam: John Benjamins

¹ Universidade Federal Minas Gerais, Belo Horizonte

² Johannes Gutenberg Universität Mainz, Gernersheim

³ Universität des Saarlandes, Saarbrücken

- Alves, F. & Gonçalves, J.L.V.R. 2003. A Relevance Theory approach to the investigation of inferential processes in translation. In: Alves 2003, 3-24
- Alves, F. & Vale, D. forthcoming. Probing the unit of translation in time: factors on the design and development of a web application for storing, annotating, and querying translation process data. In: Across Language and Cultures, vol. 2/2009, special issue edited by R. Jääskeläinen & S. Göpferich
- Carl, M., Jensen, K.T.H., & Jakobsen, A.L. 2008. Studying human translator behavior with user-activity data. In: Natural Language Processing and Cognitive Science (NLPCS 2008) conference - part of the 10th International Conference Enterprise Information Systems 12-16 June, 2008, Barcelona, Spain
- Čulo, O., Hansen-Schirra, S., Neumann, S. & Vela, M. 2008. Empirical studies on language contrast using the English-German comparable and parallel CroCo corpus. In: Proceedings of "Building and Using Comparable Corpora", LREC 2008 Workshop, Marrakesh, Morocco, 31 May 2008, 47-51
- Hansen, S. 2003. The Nature of Translated Text. An Interdisciplinary Methodology for the Investigation of the Specific Properties of Translations. Saarbrücken: Saarbrücken Dissertations in Computational Linguistics and Language Technology. Vol. 13
- Hansen-Schirra, S., Neumann, S. & Vela, M. 2006. Multi-dimensional Annotation and Alignment in an English-German Translation Corpus. In: Proceedings des Workshops Multi-dimensional Markup in Natural Language Processing (NLPXML-2006) at EACL, 4 April, 2006, Trento, Italy. 35-42
- Hansen-Schirra, S., Neumann, S. & Steiner, E. 2007. Cohesive explicitness and explicitation in an English-German translation corpus. In: Languages in Contrast 7.2: 205-229
- Pagano, A., Magalães, C. & Alves, F. (eds.) 2005. Competência em Tradução: Cognição e Discurso. Belo Horizonte: Editora UFMG
- Steiner, E. 2008. Empirical studies of translations as a mode of language contact - "explicitness" of lexicogrammatical encoding as a relevant dimension. In: Siemund, P. & Kintana, N. (eds.) Language contact and contact languages. Amsterdam: John Benjamins. 317-346