

Crossroads between Contrastive Linguistics, Translation Studies and Machine Translation

TC3 II

Edited by

Oliver Czulo

Silvia Hansen-Schirra

Translation and Multilingual Natural
Language Processing 4



Translation and Multilingual Natural Language Processing

Editors: Oliver Čulo (Universität Leipzig), Silvia Hansen-Schirra (Johannes Gutenberg-Universität Mainz), Stella Neumann (RWTH Aachen), Reinhard Rapp (Johannes Gutenberg-Universität Mainz)

In this series:

1. Fantinuoli, Claudio & Federico Zanettin (eds.). New directions in corpus-based translation studies.
2. Hansen-Schirra, Silvia & Sambor Grucza (eds.). Eyetracking and Applied Linguistics.
3. Neumann, Stella, Oliver Čulo & Silvia Hansen-Schirra (eds.). Annotation, exploitation and evaluation of parallel corpora: TC3 I.
4. Čulo, Oliver & Silvia Hansen-Schirra (eds.). Crossroads between Contrastive Linguistics, Translation Studies and Machine Translation: TC3 II.
5. Rehm, Georg, Felix Sasaki, Daniel Stein & Andreas Witt (eds.). Language technologies for a multilingual Europe: TC3 III.
6. Menzel, Katrin, Ekaterina Lapshinova-Koltunski & Kerstin Anna Kunz (eds.). New perspectives on cohesion and coherence: Implications for translation.
7. Hansen-Schirra, Silvia, Oliver Čulo, Sascha Hofmann & Bernd Meyer (eds.). Empirical modelling of translation and interpreting.
8. Svoboda, Tomáš, Łucja Biel & Krzysztof Łoboda (eds.). Quality aspects in institutional translation.

Crossroads between Contrastive Linguistics, Translation Studies and Machine Translation

TC3 II

Edited by

Oliver Czulo

Silvia Hansen-Schirra

Oliver Czulo & Silvia Hansen-Schirra (eds.). 2017. *Crossroads between Contrastive Linguistics, Translation Studies and Machine Translation: TC3 II* (Translation and Multilingual Natural Language Processing 4). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/102>

© 2017, the authors

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/>

ISBN: 978-3-946234-26-5 (Digital)

978-3-946234-98-2 (Hardcover)

ISSN: 2364-8899

DOI:10.5281/zenodo.1019701

Source code available from www.github.com/langsci/102

Collaborative reading: paperhive.org/documents/remote?type=langsci&id=102

Cover and concept of design: Ulrike Harbort

Typesetting: Sebastian Nordhoff, Iana Stefanova, Florian Stuhlmann

Proofreading: Felix Kopecky, Jean Nitzke, Valeria Quochi

Fonts: Linux Libertine, Arimo, DejaVu Sans Mono

Typesetting software: Xe_{La}TeX

Language Science Press

Unter den Linden 6

10099 Berlin, Germany

langsci-press.org

Storage and cataloguing done by FU Berlin



Language Science Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

Preface to the new edition	v
1 Introduction	
Oliver Czulo & Silvia Hansen-Schirra	1
2 Inside the monitor model: processes of default and challenged translation production	
Michael Carl & Barbara Dragsted	5
3 Text structure in a contrastive and translational perspective: On information density and clause linkage in Italian and Danish	
Lørn Korzen & Morten Gylling	31
4 Methodological cross-fertilization: Empirical methodologies in (computational) linguistics and translation studies	
Erich Steiner	65
5 An analysis of translational complexity in two text types	
Martha Thunes	91
6 Statistical machine translation support improves human adjective translation	
Gerhard Kremer, Matthias Hartung, Sebastian Padó & Stefan Riezler	121
7 Abstract pronominal anaphors and label nouns in German and English: Selected case studies and quantitative investigations	
Heike Zinsmeister, Stefanie Dipper & Melanie Seiss	153
Index	197

Preface to the new edition

This volume is a re-issue of the second of three volumes made up of previous issues of the open access journal "Translation: Computation, Corpora, Cognition" (TC3) which was transformed into the book series "Translation and Multilingual Natural Language Processing" (TMNLP) at LangSci Press. The underlying TC3 issue focused on the potential of exchange between the three fields Contrastive Linguistics, Translation Studies and Machine Translation. Today, we can look back and ask what has changed in the last five years.

The corpus paradigm proved a way to establish an intensified exchange between Translation Studies and Contrastive Linguistics, both in terms of methodology and theory. It seems, however, that the flow of information is stronger in one direction: translation scholars have adapted and evolved approaches from Contrastive Linguistics, yet our observation is that this has happened to a lesser extent in the opposite direction.

In Contrastive Linguistics, there has been a growing interest in the analysis of translations from various viewpoints, though apparently not so much in translation theory itself. Among those who integrate translations in their analysis of language contrasts is Gast (2015) who makes use of the Europarl corpus to study impersonalisation strategies in English and to contrast those to strategies found in the German data. Levshina (2017) investigates the use of *t/v* forms (formal vs. informal forms of addressing) in various languages by means of a parallel corpus of subtitles. While the design of these studies may be carefully crafted around the specificities of translated language, one should not forget the methodological drawbacks of using parallel data for contrastive studies: Translated language is influenced by the translation process and potentially by the source language. If not analysed with this in mind, translation data can be misleading, as Neumann & Hansen-Schirra (2013) show in their discussion of the differences between the translational and the contrastive perspective.

For Machine Translation and Translation Studies, the topic of post-editing was and is a point of convergence. After all, this is the point at which humans and machines meet. Typically, the topic of machine translation is addressed from angles such as quality issues, evaluation of machine translation or the impact

of machine translation and post-editing on translations and translators. These topics are much discussed - and rightly so! - in the field of post-editing. However, little work has been done so far on how to model machine translation in terms of translation-theoretic models. Exceptions are, for instance, Rozmyslowicz (2014) who discusses whether translated texts produced by a machine can be counted as translations even though functional translation theory requires agency for the production of a translation, or Lapshinova-Koltunski (2013) who treats machine-translated texts as a text type in its own right and analyses them for their linguistic variation. While the latter is not actually a contribution in terms of theoretical modelling of machine translation, it does represent an interesting shift in perspective on how machine-translated texts are treated. Čulo (2014) makes an attempt at classifying machine translation from the viewpoint of functional theories, albeit on a rather abstract level.

This re-issue contains two new paradigms besides the corpus paradigm which have gained ground in all of the fields addressed here: the cognitive and process-based research paradigm, which interact heavily. With techniques such as eye tracking and key logging, we gain insight into micro and macro processes in various types of text processing, including post-editing, which also brings Machine Translation into focus Carl et al. (see e.g. Carl & Dragsted, this volume; 2016). We hope that this re-edition will give another impulse for connecting Translation Studies with its neighbouring fields, potentially discovering new (common?) grounds.

Leipzig, Gernersheim and Aachen, September 2017
Oliver Czulo, Silvia Hansen-Schirra, Stella Neumann

References

- Carl, Michael, Srinivas Bangalore & Moritz Schaeffer. 2016. *New directions in empirical translation process research: Exploring the CRITT TPR-DB*. Cham: Springer International Publishing.
- Čulo, Oliver. 2014. Approaching Machine Translation from Translation Studies: a perspective on commonalities, potentials, differences. In *Proceedings of the Seventeenth Annual Conference of the European Association for Machine Translation (EAMT)*, 199–208. Dubrovnik, Croatia.
- Gast, Volker. 2015. On the use of translation corpora in contrastive linguistics: A case study of impersonalization in english and german. *Languages in Contrast* 15(1). 4–33.

- Lapshinova-Koltunski, Ekaterina. 2013. VARTRA: A comparable corpus for the analysis of translation variation. In *Proceedings of the 6th Workshop on Building and Using Comparable Corpora*, 77–86. Sofia, Bulgaria.
- Levshina, Natalia. 2017. A multivariate study of t/v forms in European languages based on a parallel corpus of film subtitles. *Research in Language* 15(2). 153–172.
- Neumann, Stella & Silvia Hansen-Schirra. 2013. Exploiting the incomparability of comparable corpora for Contrastive Linguistics and Translation Studies. In Serge Sharoff, Reinhard Rapp, Pierre Zweigenbaum & Pascale Fung (eds.), *Building and Using Comparable Corpora*, 321–335. Berlin: Springer.
- Rozmyslowicz, Tomasz. 2014. Machine Translation: A Problem for Translation Theory. *New voices in translation studies* 10(11). 145–163.

