

Empirical studies in translation and discourse

Edited by

Mario Bisiada

Translation and Multilingual Natural
Language Processing 14



Translation and Multilingual Natural Language Processing

Editors: Oliver Czulo (Universität Leipzig), Silvia Hansen-Schirra (Johannes Gutenberg-Universität Mainz), 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. Czulo, 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 Czulo & Sascha Hofmann (eds.). Empirical modelling of translation and interpreting.
8. Svoboda, Tomáš, Łucja Biel & Krzysztof Łoboda (eds.). Quality aspects in institutional translation.
9. Fox, Wendy. Can integrated titles improve the viewing experience? Investigating the impact of subtitling on the reception and enjoyment of film using eye tracking and questionnaire data.
10. Moran, Steven & Michael Cysouw. The Unicode cookbook for linguists: Managing writing systems using orthography profiles.
11. Fantinuoli, Claudio (ed.). Interpreting and technology.
12. Nitzke, Jean. Problem solving activities in post-editing and translation from scratch: A multi-method study.
13. Vandevoorde, Lore. Semantic differences in translation.
14. Bisiada, Mario (ed.). Empirical studies in translation and discourse.

Empirical studies in translation and discourse

Edited by

Mario Bisiada

Bisiada, Mario (ed.). 2021. *Empirical studies in translation and discourse* (Translation and Multilingual Natural Language Processing 14). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/296>

© 2021, the authors

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

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

ISBN: 978-3-96110-300-3 (Digital)

978-3-96110-301-0 (Hardcover)

ISSN: 2364-8899

DOI: 10.5281/zenodo.4450014

Source code available from www.github.com/langsci/296

Collaborative reading: paperhive.org/documents/remote?type=langsci&id=296

Cover and concept of design: Ulrike Harbort

Typesetting: Mario Bisiada

Proofreading: Amir Ghorbanpour, Anca Găță, Dora Uștulica, Elen Le Foll,

Jeroen van de Weijer, Lachlan Mackenzie, Marten Stelling, Jean Nitzke, Tihomir

Rangelov, Tom Bossuyt

Fonts: Libertinus, Arimo, DejaVu Sans Mono

Typesetting software: \LaTeX

Language Science Press

xHain

Grünberger Str. 16

10243 Berlin, Germany

langsci-press.org

Storage and cataloguing done by FU Berlin

Freie Universität  Berlin

Contents

Preface	iii
1 Post-editing: a genealogical perspective on translation practice Maeve Olohan	1
2 Testing the Gravitational Pull Hypothesis on modal verbs expressing obligation and necessity in Catalan through the COVALT corpus Josep Marco	27
3 Stylistic normalisation, convergence and cross-linguistic interference in translation: The case of the Czech transgressive Olga Nádvorníková	53
4 The internationalized text and its localized variations: A parallel analysis of blurbs localized from English into Arabic and French Madiha Kassawat	93
5 Movement or debate? How #MeToo is framed differently in English, Spanish and German Twitter discourse Mario Bisiada	113
6 Investigating patterns of saccadic eye movement when using Microsoft's Skype Translator between Catalan and German Felix Hoberg	141
7 What can Euclidean distance do for translation evaluations? Éric André Poirier	165
8 Between audiovisual translation and localization: The case of <i>Detroit: Become Human</i> Laura Mejías-Climent	199
9 Analysing the dimension of mode in translation Ekaterina Lapshinova-Koltunski	221
Index	242

Preface

Empirical translation studies has moved beyond being a mere methodological approach within translation studies to becoming an established sub-field of translation studies. This is not only shown by the amount of volumes dedicated to it (Hansen 2002; Carl et al. 2015; Ji 2016; de Sutter et al. 2017; Ji & Oakes 2019). The idea of empirical translation studies has a strong tradition, and perhaps to some extent its origin, in corpus-based translation studies (de Sutter et al. 2012; Ji et al. 2016), but empirical approaches have also been adopted in other paradigms of translation studies (Laviosa et al. 2016).

In their programmatic article *Towards methodologically more rigorous corpus-based translation studies*, de Sutter et al. (2012) offer a set of minimal requirements for research papers in the field of corpus-based translation studies. According to this, papers should:

1. “provide a meticulous overview of the corpus materials used and of the exact procedures for selecting, annotating and sifting the data”
2. “comment on any specific problems encountered during data selection and annotation, including explicit and motivated statements as to the solutions being adopted”
3. “include elaborate testing for statistical significance as a complement of, not in opposition to, thorough qualitative analysis”

With the slight modification that the data does not necessarily have to be corpus data, I think we can adopt at least the first two requirements for a wider view of empirical translation studies, as those are steps that allow other researchers to reproduce the study in question. As for the issue of statistical significance, it is true that the uptake has been slow in our discipline, though studies generally tend to include significance testing. At the same time, the notion of “statistical significance” is continuously questioned in science in general (see, e.g., McShane et al. 2019) and in corpus linguistics in particular (Koplenig 2019).

A more recent suggestion for an improved research agenda in empirical translation studies comes from de Sutter & Lefer (2019), who demand passing on from

the analysis of well-studied phenomena towards exploring new linguistic indicators, such as “linguistic features that have been said to typify other forms of constrained communication, such as non-native language varieties, editing and student writing” (de Sutter & Lefer 2019: 19). Starting from the basic assumption that “translation products and processes are multifaceted and multidimensional” (de Sutter & Lefer 2019: 18), so that their study should draw on multi-methodological designs. They argue that “understanding translation inevitably entails an interdisciplinary approach to translation, building on theoretical frameworks and findings from neighbouring disciplines, including, but not restricted to, variational corpus linguistics, bilingualism studies and (cognitive) sociolinguistics” (de Sutter & Lefer 2019: 18).

It is in this spirit, then, that the present volume seeks to contribute some studies to the subfield of Empirical Translation Studies and thus aid in extending its reach within the field of translation studies and thus in making our discipline more rigorous and fostering a reproducible research culture. The Translation in Transition conference series, across its editions in Copenhagen (2013), Garmersheim (2015) and Ghent (2017), has been a major meeting point for scholars working with these aims in mind, and the conference in Barcelona (2019) has continued this tradition of expanding the sub-field of empirical translation studies to other paradigms within translation studies. This book is a collection of selected papers presented at that fourth Translation in Transition conference, held at the Universitat Pompeu Fabra in Barcelona on 19–20 September 2019.

While maintaining the interdisciplinary focus and the strong standing of corpus-aided research in empirical translation studies, the conference has also seen input from fields such as audiovisual translation studies, cross-linguistic discourse studies. The papers in this volume are ordered roughly by the type of language they analyse, on a cline from primarily written to primarily spoken language, passing on its way via a hybrid type of conversational or “informally written” (McCulloch 2019) language. A connecting thread is the analysis of features and style of language in multilingual discourse environments. This can be seen in the chapters by Josep Marco, Éric Poirier, and Ekaterina Lapshinova-Koltunski chapters, which suggest ways of analysing the shifts of informationally and linguistically salient elements in the target texts as well as in Olga Nádvořníková’s, Laura Mejías-Climent’s and Madiha Kassawat’s chapters analysing stylistic features of the target language in entertainment products, as well as in Bisiada’s chapter, which compares the semantic features of particular expressions across languages to represent feminist movements.

Maeve Olohan develops a practice-theoretical conceptualisation of post-editing as one of several activities that make up the changing practice of translating,

alongside other activities such as editing translation memory fuzzy matches. This contrasts with a view of post-editing as a practice in its own right that competes with or complements the practice of translating. She examines how post-editing is reconfiguring translation practice, through changes in constituent elements of the practice, including the tools and materials deployed, the competences and knowing that transpire in practice, and various understandings of the practice. By exploring this reconfiguration, we may extend our genealogical understanding of translation, as a practice that changes over time.

Josep Marco examines Halverson's (2010) Gravitational Pull Hypothesis which draws on cognitive linguistics and bilingual theory to address the problem that translated texts have been shown in some cases to over-, in others to under-represent of typical target language elements, which creates the dilemma that both claims cannot be generally true or predicated of the same set of data. Halverson (2010) argues that patterns of prototypicality in the target language, conceptual structures or the representation of the source language item, and patterns of connectivity are possible cognitive causes of this issue. Marco draws on the COVALT corpus consisting of English-Catalan and French-Catalan parallel corpora and a Catalan comparable corpus to test her hypothesis on the Catalan verb *caldre*.

Olga Nádvoříková argues in her chapter that the Czech transgressive is a non-finite verb form belonging to the cross-linguistic converb category. In contrast with other converbs, the Czech transgressive has a strong stylistic mark and is very rare in contemporary language. Using a parallel multilingual corpus and a comparable corpus of translated and non-translated Czech, the chapter investigates the differences in the frequency of the transgressive in translated and non-translated fiction and non-fiction. The data shows the effect of convergence in both fiction and non-fiction and the effect of stylistic normalisation in fiction only. The results of the potential effect of cross-linguistic interference are much less conclusive, indicating that a thorough contrastive analysis of different language pairs is required first.

Madiha Kassawat argues in her chapter that, in an increasingly globalised world, accessibility to digital content has become indispensable for people around the world, which in turn makes translation indispensable. As the majority of products is promoted for and sold on the internet, their web pages are often localised based on the market. The required speed in this type of work, its tools and process influence the quality of the localised texts, which necessitates their analysis and an exploration of the different interpretations of the same source text in several languages. Her study compares the product descriptions provided in English and localised into Arabic and several French versions.

Mario Bisiada studies tweets on #MeToo in English, Spanish and German from 2019, revealing how MeToo is most commonly referred to as a “movement” in English and Spanish but as a “debate” in German, a difference that echoes German-language press habits. Based on an analysis of semantic prosody, the chapter demonstrates that words indicating longevity such as *era* and *times* collocate with MeToo in English and Spanish, but not in German. This points to a framing of MeToo as influential and long-term in English and Spanish and as exaggerated and short-term in German. Reflecting this difference, MeToo is talked about in more negative terms in German tweets compared to English and Spanish, as shown by a qualitative analysis of evaluative author stance.

Felix Hoberg investigates the patterns of saccadic eye movement when using Microsoft’s Skype Translator between Catalan and German through a case study of 21 German-speaking participants as part of an overall evaluation of the Skype Translator on a dialogue-oriented level. Despite not having any proficiency in Catalan, these participants had to text-chat with Catalan native speakers via Skype, while the Skype Translator was activated. The sessions were observed by an eye tracking system. The collected data thus represents a naturalistic starting point to evaluate how users structure computer-mediated communication situations when real-time machine translation is involved while having to rely on that output.

Éric Poirier describes an empirical method to screen informational translation shifts in parallel segment pairs extracted from bilingual or multilingual translation corpora, based on character length and lexical word count. The method applies to most known languages and in one or the other of the two translation directions (direct or inverse). The chapter argues that heteromorphic segment pairs, as opposed to isomorphic ones, are more likely to contain informational translation shifts. The objective and reproducible method described in his chapter allows for semi-automatic identification of problematic translations and uncovering of textual and linguistic facts revealing translation processes, contingencies, and determinism.

Laura Mejías-Climent’s chapter is an analysis of the dubbing of the video game *Detroit: Become Human*. She wants to shed some light on the convergences of audiovisual translation studies and localisation from the specific perspective of dubbing, in a product that, in turn, poses some questions to the genre it belongs to. This chapter aims to highlight some of the differences and convergences between AVT and localisation analysing the dubbing synchronies applied in a video game belonging to a genre closer to traditional movies, compared to other adventure games, due to the strong presence of cinematic scenes and the lower level of interaction.

Ekaterina Lapshinova-Koltunski analyses English-to-German translations and interpretations, focussing on the variation in English-to-German translation that involves the dimension of mode, i.e. variation between spoken and written language production. She argues that the resulting variation is reflected in the linguistic features of translations and interpretations, e.g. preferences for modality meanings, proportion of nominal or verbal phrases and others. These features offer the opportunity of analysing and modelling the dimensions involved. The methodological focus of her chapter is on quantitative distributions of these linguistic features reflected in the lexico-grammar of texts.

As is evident, the studies are united by a strong empirical aspect, based for instance on corpus analyses or eye-tracking approaches, and a few also have a theoretical focus on features of translated language or the effect of post-editing on translation practice at the workplace. The studies come from subfields as diverse as audiovisual translation, machine translation, cultural mediation and contrastive linguistics and include a range of languages such as Spanish, German, Arabic, Czech, Catalan, French as well as English.

This volume was produced in the difficult circumstances of a pandemic, and I would like to thank everyone involved for their efforts so this volume could appear at the time originally envisaged after the conference. Thanks go to the peer reviewers Miquel Pujol Tubau, Simon Varga, Sandra Halverson, Silvia Hansen-Schirra, Václav Cvrček, Jennifer Fest, Dorothy Kenny, Oliver Czulo and those who wish to stay anonymous, the proofreaders and the team at Language Science Press, for their timely work and for believing in open science.

Mario Bisiada
Barcelona, January 2021

References

- Carl, Michael, Srinivas Bangalore & Moritz Schaeffer (eds.). 2015. *New directions in empirical translation process research: Exploring the CRITT TPR-DB*. Heidelberg: Springer.
- de Sutter, Gert, Patrick Goethals, Torsten Leuschner & Sonia Vandepitte. 2012. Towards methodologically more rigorous corpus-based translation studies. *Across Languages and Cultures* 13(2). 137–143.
- de Sutter, Gert & Marie-Aude Lefer. 2019. On the need for a new research agenda for corpus-based translation studies: A multi-methodological, multifactorial and interdisciplinary approach. *Perspectives: Studies in Translation Theory and Practice* 28(1). 1–23.

- de Sutter, Gert, Marie-Aude Lefer & Isabelle Delaere. 2017. *Empirical translation studies: New methodological and theoretical traditions*. Berlin: de Gruyter.
- Halverson, Sandra. 2010. Cognitive translation studies: Developments in theory and method. In Gregory M. Shreve & Erik Angelone (eds.), *Translation and cognition*, 349–70. Amsterdam: John Benjamins.
- Hansen, Gyde. 2002. *Empirical translation studies: Process and product*. Copenhagen: Samfundslitteratur.
- Ji, Meng. 2016. *Empirical translation studies: Interdisciplinary methodologies explored*. London: Equinox.
- Ji, Meng, Lidun Hareide, Defeng Li & Michael Oakes (eds.). 2016. *Corpus methodologies explained: An empirical approach to translation studies*. London: Routledge.
- Ji, Meng & Michael Oakes (eds.). 2019. *Advances in empirical translation studies: Developing translation resources and technologies*. Cambridge: Cambridge University Press.
- Koplenig, Alexander. 2019. Against statistical significance testing in corpus linguistics. *Corpus Linguistics and Linguistic Theory* 15(2). 321–346.
- Laviosa, Sara, Adriana Pagano, Hannu Kemppanen & Meng Ji. 2016. *Textual and contextual analysis in empirical translation studies*. Heidelberg: Springer.
- McCulloch, Gretchen. 2019. *Because internet: Understanding how language is changing*. London: Random House.
- McShane, Blakeley B., David Gal, Andrew Gelman, Christian Robert & Jennifer L. Tackett. 2019. Abandon statistical significance. *The American Statistician* 73. 235–245.