

CTT 2024

**The 1st Workshop on Creative-text Translation and
Technology**

Proceedings

June 27, 2024

The CTT organizers gratefully acknowledge the support from the following sponsors.

With the support of





The papers published in this proceedings are —unless indicated otherwise— covered by the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC-BY-NCND 4.0). You may copy, distribute, and transmit the work, provided that you attribute it (authorship, proceedings, publisher) in the manner specified by the author(s) or licensor(s), and that you do not use it for commercial purposes. The full text of the licence may be found at <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>

©2024 The authors

ISBN 978-1-0686907-3-0

Publisher: European Association for Machine Translation (EAMT)

Preface by the Workshop Organizers

This volume contains the contributions to the first workshop on Creative-text Translation and Technology (CTT).¹ CTT is co-located with the 25th Annual Conference of the European Association for Machine Translation (EAMT 2024)², held on 27 June 2024 in Sheffield, UK.

Scope. In an era where technological advances continuously reshape the potential of language tools, the call for papers asked for novel work that discusses the creative aspects of language technology. Our scope was therefore deliberately broad, envisioning work on the use of language technology for the translation of texts in creative domains, such as marketing, literature and poetry, audiovisual translation, and multi-lingual content creation. Tools could include large language models (LLMs), computer-aided translation (CAT) tools, machine translation (MT) systems, post-editing environments, and so on. Furthermore, we aimed to attract submissions from a diverse range of profiles: researchers, educators, translators and industry stakeholders were all encouraged to submit their work to ensure a broad platform for discussion focused on the suitability of current language technology for different creative translation processes.

Submissions. We received 9 submissions in total. After a double-blind peer-review process, two papers were rejected because their focus on language technology was too limited in scope. We encouraged those authors to submit their work to more topical venues. With 7 out of 9 papers ultimately accepted, we have an acceptance rate of 78%, with an average reviewer score of 16/20. All papers are between 5 and 9 pages long and accepted for oral presentation.

Given the current technological landscape, it does not come as a surprise that LLMs emerged as a topic of significant interest. Van Egdom, Declercq, & Kusters investigate how post-editing through prompting can improve the quality of machine-translated literary work. Macken presents a study that compares the quality of professional (human) post-editing and post-editing with ChatGPT in a selection of short stories. And Castaldo & Monti focus on a specific type of creative language translation, namely idioms, and analyse the quality of LLMs when translating these expressions.

With regards to workflows, Daems, Ruffo, & Macken compare the effect of using different types of translation workspaces (Word, Trados, and a literary post-editing platform) on translation text features, such as sentence length and style, in the context of short story translation. Oliver & Alvarez-Vidal report on the LitPC toolkit, a set of tools that can aid in building parallel corpora from literary works, which in turn can be used to train machine translation systems. Mikelenić & Oliver introduce neural MT systems tailored to literature in Spanish, French, Italian, Portuguese and Romanian, leveraging both multilingual literary corpora and web-crawled datasets. And Jon & Bojar investigate surprisal distributions in the source text on the one hand and MT and human translation on the other, hypothesising that MT models enforce a uniform surprisal distribution.

Keynotes. We have the pleasure to host two keynote speakers at CTT.³ Ana Guerberof Arenas will present the work done on creativity in machine translation and beyond in her past project CREAMT and her newly-acquired ERC project INCREC. Andrew Rothwell will discuss how literary translators can “augment” their creativity in this fast-changing technological landscape while providing an overview of the various technological tools that are available.

Sponsors. CTT is grateful to be sponsored by INTERACT: Interdisciplinary research network on language contact research⁴, funded by the Research Foundation Flanders with grant number W002220N.

¹<https://ctt2024.ccl.kuleuven.be/>

²<https://eamt2024.sheffield.ac.uk/>

³<https://ctt2024.ccl.kuleuven.be/keynotes>

⁴<https://interact.ugent.be/>

CTT is also sponsored by the research group on Computational and Formal Linguistics (ComForT)⁵ at KU Leuven.

With this first workshop on Creative-text Translation and Technology, we aim to bring together a diverse audience to talk about the applicability of language technology to creative-text translation. We are looking forward to the fruitful discussions and insights.

June 2024,

Bram Vanroy, Marie-Aude Lefer, Lieve Macken, Paola Ruffo

⁵<https://www.arts.kuleuven.be/ling/comfort-english/>

Organizing Committee

Program Chairs

Bram Vanroy, Instituut voor de Nederlandse Taal and KU Leuven

Marie-Aude Lefer, UCLouvain

Lieve Macken, Universiteit Gent

Paola Ruffo, Universiteit Gent

Program Committee

Reviewers

Lynne Bowker, University of Ottawa
Joke Daems, Universiteit Gent
Ana Guerberof-Arenas, University of Groningen
James Luke Hadley, University of Dublin, Trinity College
Damien Hansen, University of Liège
Alina Karakanta, Leiden University
Dorothy Kenny, Dublin City University
Waltraud Kolb, University of Vienna
Maarit Koponen, University of Eastern Finland
Antoni Oliver, Universitat Oberta de Catalunya
Paola Ruffo, Universiteit Gent
Minna Ruokonen, University of Eastern Finland
Kristiina Taivalkoski-Shilov, University of Turku
Arda Tezcan, Universiteit Gent
Antonio Toral, University of Groningen
Susana Valdez, Universiteit Leiden
Chantal Wright, ZHAW - Zürcher Hochschule für Angewandte Wissenschaften
Mehmet Şahin, Boğaziçi University

Keynote Talk

The INCREC project: creativity and technology in translation

Ana Guerberof Arenas
University of Groningen, The Netherlands



Thu, June 27, 2024 – Time: 09:00 – 10:00 –

Abstract: In this talk, I will give an overview of how creativity is conceptualised in the social sciences, mainly psychology and sociology, including different frameworks that facilitate analysing creativity. I will also touch upon how the technological field has presented and has studied creativity, and how translation in combination with technology can be explored.

With this aim in mind, I will also present results from the CREAMT project (2020-2022) that explored creativity in literary texts in different translation modalities: translation by professional literary translators, machine translation using a customized neural engine, and post-edition. Further it looked at the impact on readers by looking at narrative engagement, enjoyment and translation reception.

Finally, I will describe the new ERC CoG INCREC research project (2023-2028) that looks to uncover the creative process in professional literary and audiovisual translators in order to create specific frameworks, and how and when technology, e.g. machine translation, can be used to enhance rather than constrain creativity. But also if the intended audiences, readers and viewers, appreciate, not only cognitively but also emotionally, creative shifts in translated content and why this might be.

Bio: Ana Guerberof Arenas is an associate professor in Translation Studies at University of Groningen. From 2020-22, she was a Marie Skłodowska Curie Research Fellow working on the CREAMT project that looked at the impact of MT on translation creativity and the reader's experience in the context of literary texts. More recently she has been awarded a ERC Consolidator grant by the European Research Council to work on the five-year project INCREC (2023-2028) that explores the translation creative process in its intersect with technology in literary and AVT translation. She has authored refereed articles and book chapters on MT post-editing; reading comprehension of MT output; translator training, ethical considerations in MT, AI and the industry, creativity and reception studies. She has more than 23 years' experience in the translation industry.

Keynote Talk

CAT, TM, NMT, and AI: A Literary Translator's Dream Team?

Andrew Rothwell
Swansea University, United Kingdom



Thu, June 27, 2024 – Time: 13:30 – 14:30 –

Abstract: Since the public release of ChatGPT in November 2022, there has been an explosion of interest in the remarkable text-production capabilities (including paraphrasing, summarizing and translation) of such generative AI tools. Whether they can be of assistance to the literary translator, and if so, how they can best be made to interoperate with existing CAT and MT environments, remains, however, a largely moot question.

This paper will describe my developing use of different technologies, over almost a decade, to produce English translations of classic novels in French by Emile Zola and Marcel Proust. Acknowledging the documented reticence of literary translators to adopt computerised tools, I will nevertheless argue for the practical benefits of using: –

- an electronic ST
- aligned bilingual editor (aka CAT tool)
- translation memory
- termbase
- online dictionaries and thesauri
- NMT (free-standing and CAT-integrated)
- generative AI.

The core of the paper will be a presentation of how these technologies are now being combined in a single interface, taking as an example the recently AI-enhanced CAT tool Wordscope. Wordscope offers an integration of translation memory, machine translation from several providers, and ChatGPT as a research and paraphrasing tool, in a de-cluttered online environment.

The paper will describe different options for using the tool for literary translation, and discuss some theoretical implications of doing so in a Cognitive Translation Studies framework. In conclusion, I will argue

that Lommel's (2018) notion that the translator's creative capacity is 'augmented' rather than inhibited by computer technologies applies no less to literary than to 'commercial' translation, albeit in significantly different ways.

Bio: Andrew Rothwell is Emeritus Professor of French and Translation Studies at Swansea University. He has research interests in contemporary French literature, especially the poet and art writer Bernard Noël, translation technologies, and translation theory. He has published numerous literary translations into English, including of poetry by Noël and two novels by Emile Zola (for Oxford World's Classics), and is currently working on a new translation of Marcel Proust's *La Prisonnière* for the same publisher. Recent books include *Translation Tools and Technologies* (with Joss Moorkens, María Fernández Parra, Joanna Drugan and Frank Austermuehl) and *Computer-Assisted Literary Translation* (co-edited with Andy Way and Roy Youdale), both published in 2023 by Routledge.

Table of Contents

<i>Using a multilingual literary parallel corpus to train NMT systems</i> Bojana Mikelenić and Antoni Oliver	1
<i>'Can make mistakes'. Prompting ChatGPT to Enhance Literary MT output</i> Gys-Walt Van Egdom, Christophe Declercq and Onno Kusters	10
<i>LitPC: A set of tools for building parallel corpora from literary works</i> Antoni Oliver and Sergi Alvarez-Vidal	21
<i>Prompting Large Language Models for Idiomatic Translation</i> Antonio Castaldo and Johanna Monti	32
<i>An Analysis of Surprisal Uniformity in Machine and Human Translations</i> Josef Jon and Ondřej Bojar	40
<i>Impact of translation workflows with and without MT on textual characteristics in literary translation</i> Joke Daems, Paola Ruffo and Lieve Macken	57
<i>Machine Translation Meets Large Language Models: Evaluating ChatGPT's Ability to Automatically Post-Edit Literary Texts</i> Lieve Macken	65