

Translation Research Projects 6. Testing Borders: A Tribute to Anthony Pym



Edited by

Kayoko Takeda, Ester Torres-Simón, David Orrego-Carmona,
Nune Ayvazyan and Bei Hu

Intercultural Studies Group
Universitat Rovira i Virgili

Translation Research Projects 6.
Testing Borders: A Tribute to Anthony Pym

*Edited by Kayoko Takeda, Ester Torres-Simón, David Orrego-Carmona,
Nune Aivazyan and Bei Hu.*

Tarragona, 2026

Title: Translation Research Projects 6. Testing Borders: A Tribute to Anthony Pym

Editors: Kayoko Takeda, Ester Torres-Simón, David Orrego-Carmona, Nune Ayvazyan, Bei Hu

© for texts, individual authors 2026

© for compilation, the editors, 2026

Publisher: Intercultural Studies Group. URV. Av. Catalunya 35. 43002 Tarragona, Spain.
<http://isg.urv.es/>

Printer: Printcolor

ISBN – 13: 978-84-09-88237-3

Texts available at: <https://www.intercultural.urv.cat/en/publications/translation-research-projects-6/>

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means without prior written permission from the publisher.

Front cover: Collage with pictures of Anthony Pym's students in various events.

Printed in Spain



Contents

David Orrego-Carmona, Kayoko Takeda, Ester Torres-Simón, Nune Ayvazyan and Bei Hu, <i>Introduction</i>	11
Andrew Chesterman, <i>Prologue: Anthony Pym, on borders</i>	17
Nataša Pavlović, <i>The relevance of translation theory for translator education today</i>	21
Volga Yilmaz-Gümüř, <i>Tracing the evolution of translator training in Anthony Pym’s research</i>	41
Judith Raigal-Aran and Nune Ayvazyan, <i>Watching the process unfold: using API-based interaction data to understand student use of AI in translation tasks</i>	59
Maggie Hui, <i>A constraints-first approach to CAT and generative AI in translator education: a simulated client project in a master’s course</i>	79
Enes Ekici And řeyda Eraslan, <i>Beyond perceptual failure: cognitive load and number renditions in technology-mediated English–Turkish simultaneous interpreting</i>	101
Yu Hao, <i>Quality in multilingual crisis communication: a reception-oriented perspective</i>	121
Juerong Qiu, <i>How does multimodal creativity travel? A reception perspective on cyberdubbing</i>	139
Gabriel González Núñez, <i>“Art, esthetics, ideal”: an ethics for literary translators in the age of translating machines</i>	159
Snježana Veselica Majhut, <i>How to find Paula D.? In quest of translators’ identities in early 20th century Croatia</i>	179
Kayoko Takeda, <i>On interpreting v. translation</i>	197

Introduction

A homage volume usually opens by mapping the territory its honoree made their own: the editors name a subfield, sketch its central questions, and show how one scholar redefined it. We found that move unusually hard to make for Anthony Pym, and the difficulty says something true about him.

Pym never settled in a single province of the discipline. He has worked on translation history and on how to write it, on the training of translators and the competences they need, on the ethics of the translator, on the sociology of translators as agents, on localization and machine translation, on risk and uncertainty and on the shape of translation theory itself. Alongside all of this, he kept asking where the borders of the field should be drawn, and whether there should be any borders at all. A scholar who treats the internal frontiers of his own discipline as open doors is hard to place in a particular territory.

This is why the conventional opening essay would misrepresent him. What holds his work together operates beneath the level of subject matter; it's a consistent way of going about research and knowledge. Pym asks for evidence, he reads closely and widely, he argues openly in his publications, and he keeps his eye on the people who translate and those who engage with translations. Those habits travel with him from one subfield of translation studies to the next, which is why he could move among so many without ever seeming to change discipline. This introduction, therefore, takes a different route. It follows the habits across several subareas of translation studies at once and lets the chapters show where they lead.

The book is almost entirely the work of Anthony's former students. Apart from Andrew Chesterman, who writes the foreword as a peer and old friend, every author here studied with Pym, most of them as doctoral students in Tarragona, in the Translation and Intercultural Studies program that he founded almost three decades ago. We now teach and research in Spain, Croatia, Türkiye, Hong Kong, Japan, Australia, Singapore, the United Kingdom and the United States.¹ There is no better way to assess the impact of a teacher than to see the impact on his students.

The volume also comes home in another sense. It appears in Translation Research Projects, the series Pym founded and edited at the Intercultural

¹ One of Anthony's former doctoral students, Mehrnaz Pirouznik, had hoped to contribute. A prolonged internet blackout in Iran during the 2026 war cut off her access for more than eighty days and made it impossible for her to submit in time. We note her absence here with warmth, and look forward to seeing her work in the future.

Studies Group, which, from the start, published its volumes online for anyone to download in open access, years before the term was in common use and well before funders began to require it. It seems right to give his students' work back to him in the series, and in the open form he chose for it, because that is how he has treated scholarship: as something to be shared freely, and as work that ought to reach the people it concerns.

Pym came into translation studies from outside it, by way of a doctorate in sociology, as well as experience in language teaching and professional translation. That combination marks everything he has done. His working life has been as mobile as his subject. For the most part, it has been split between Spain (La Laguna, the Autònoma de Barcelona and the Rovira i Virgili in Tarragona) and universities on five continents, among them Monterey, Stellenbosch, Beijing and Melbourne, with collaborations reaching to Ca' Foscari in Venice, Vienna, Singapore and others. He is drawn at once to the practical problems of translating and to the people who teach or do the translating. Several threads from that work run through this book. From his first monograph, *Translation and Text Transfer* (1992), he considered translation a form of intercultural communication and the translator as someone who works in the spaces where cultures overlap. In *Method in Translation History* (1998), recently fully revisited and published with Routledge (2026), he argued that the history of translation should be a history of translators as people. On training and competence, he has been steadily skeptical of elaborate models and attentive to what students actually need. His two books on ethics, beginning with *Pour une éthique du traducteur* (1997), shifted the question from the qualities of a translation to the responsibilities of the translator, and put cooperation and trust at its center. His recent work presses on uncertainty, risk and automation, and asks what a human translator is for once a machine can produce a plausible draft. That last question runs through much of what follows.

Behind all of it sits the stance Chesterman describes in the foreword: a defense of evidence and empirically testable claims, held together with a genuine willingness to say, "I don't know." Pym argues and counter-argues, points to gaps in the evidence, and invites others to do the same to him. Chesterman portrays Pym as a writer on borders, someone who has inhabited intercultural borderlands across Australia, Germany, France and Spain, who brings personal experience into academic prose as though the line between the two was never quite impervious, and who reads across esthetics, history, politics and sociology as if those borders, too, were there to be crossed. We borrow that image to open this collection, because it names something essential both about how Pym writes and about what he studies. Translation, on his account, happens in the overlaps between cultures—in intercultures—and it is sustained by people willing to negotiate, to cooperate and to take responsibility for moving meaning across a line.

The contributions gathered here aim to extend Pym's work. They share a set of concerns that map onto the concepts he has done most to put into circulation: risk and risk management (applied in Hui), the relevance of translation theory (as discussed by Pavlović), translator training (meta analyzed in Yilmaz-Gümüş), ethics (linked to literature in the article by González Núñez and reflecting on interpreting in Takeda), cooperation and trust (linked to technologies in Raigal-Aran and Ayvazyan), critical integration of technologies (see Ekici and Eraslan on interpreting and Hao on emergency translation) and the idea of interculturalities (present in Qiu, but also in Majhut). What gives the volume its coherence, beyond the figure it honors, is a shared question that several chapters pose in different registers: in an age of capable translating machines, what is the human translator's distinctive contribution, and how should we train, judge and think about it?

That question turns out to be a deeply Pymian one. It is answered here in the way he taught us to work: through classroom data, reception studies, close textual reading, archival reconstruction and careful arguments.

Thus, Nataša Pavlović's chapter investigates the empirical relevance of translation theory in translator education, especially when technological change, generative AI, platform labor and employability pressures are revolutionizing both the translator profession and the humanities. Drawing on student responses to a flipped translation theory course at the University of Zagreb, the chapter argues that translation theory remains valuable when it is connected to practice, critical reflection and professional self-understanding. It echoes Pym's longstanding proposition regarding the empirical use of theory. As Pavlović's study shows, theory can work as a mode of trial-and-error testing and reflexive inquiry, helping translators explain their choices and trade-offs and hopefully reach a more ethical judgment.

Still linked to theory, Volga Yilmaz-Gümüş sets out to provide a meta-look at Pym's contribution to translator training research based on a systematic reading of fifty-four works published between the 1990s and 2025. Her review divides Pym's contributions into five principal topics: distance learning, transferable skills, technology in translator training, employability and other areas. This last section encompasses several real-life empirical studies in translator training, a focus on less discussed areas, like internships and exchange programs and the active dialogue between language education and practicing translators that Pym encouraged. This chapter summarizes key outcomes that can offer guidance on how to approach translator training today and tomorrow. Answering this call for integrating technologies into pedagogy in ways that are both exploratory and critically informed, Judith Raigal-Aran and Nune Ayvazyan report on an experiment designed to move beyond self-report studies towards the systematic analysis of interaction logs. A 90-minute translation task done by students was tracked by an API and reported on by the students. The methodological approach allows for a comparison between

what students report and what they actually do, opening new avenues to a critical reflection on technological use.

Maggie Hui draws explicitly on Pym's emphasis on risk management as a foundational translator competence to engage in a single-case exploratory study on a student group that must carry out a technical translation from English into Chinese using a GenAI-supported computer-assisted translation (CAT) tool. By documenting specific risks in the workflow, such as additions or terminological drifts, in a constraint-aware classroom, this empirical study makes a strong case for training that puts risk management ahead of tool enthusiasm.

While often researched within translation, technologies also affect interpreting. Against the background of the increasing use of automatic speech recognition (ASR) tools in professional interpreting, Enes Ekici and Şeyda Eraslan examine whether students can take advantage of such tools to improve their numerical rendition accuracy in English-Turkish simultaneous interpreting. The results of an experiment of students interpreting number-dense speeches with and without ASR-simulated captions are analyzed through Gile's Effort Model and Seeber's Cognitive Load Model as well as Pym's risk management framework.

With a critical approach to a realistic use of technologies, Yu Hao's chapter revisits what counts as a "good" translation in multilingual crisis communication, in which translated messages must be timely, understandable, trustworthy and actionable. The chapter argues for a reception-oriented perspective grounded in the experiences of actual recipients, especially culturally and linguistically diverse communities, during health crises and disasters. It offers a tribute to Pym's intellectual influence *Down Under* over the past decade and offers a forward-looking account of translation effects in an age of crisis and uncertainty.

Building on Pym's (2020) insight that new forms of translation, such as adaptation, rewriting and transcreation and drafting, and bridging technologies and sociological approaches to translation, Qiu investigates the phenomenon of cyberdubbers. Moving away from discussions on quality or ethics, Qiu argues for a cooperative interpretation of authenticity built on the interactions between cyberdubbers and audiences creating in the process an alternative audiovisual translation product.

Gabriel González Núñez reflects on the ethics of literary translation in an age increasingly oversaturated and affected by machine translation and generative AI. Drawing on Pedro Figari's philosophy of aesthetics, the chapter argues that literary translation remains a human activity insofar as it concerns emotional appeal, esthetic responses and an ethical aspiration to better the human condition. Applying Figari's ethics of the ideal to literary translation, González Núñez critically engages with Pym's scholarship on translator ethics. Focusing on readers, this chapter attempts to extend Pym's

cooperation-based ethics to human agency and ethical decision-making in literary translation.

Following Pym's call for focusing on translators in the study of translation history, Snježana Veselica Majhut examines how translators are recognized in the peritext of "Zabavna biblioteka", a publishing project that continuously released translated fiction into Croatian from 1913 to 1941. While motivations for concealing or revealing translators' identities are analyzed, an in-depth inquiry into three translators' biographical data indicates the fluid professional profiles of those who engaged in translation in this project.

Last, Takeda looks back on how her academic horizons were broadened through the doctoral program in Tarragona, including the fact that conversations with Pym inspired her pursuit of characteristics unique to interpreting as opposed to written translation. Referring to Pym's discussion of immediate feedback in face-to-face communication as well as translator ethics, Takeda presents the risks interpreters can face in the context of war crimes. Given the ever-diversifying methods of interlingual communication, however, Takeda concludes with her shifting attention to "human v. AI" in interpreting, which is directly connected to Pym's recent works and most of the other contributions in this volume.

At its core, this volume shows that Pym's questions outlast any single tool or technique. The machines will change again, and these chapters will date; the habits of mind behind them, evidence, close attention, a refusal to look away from the person doing the work, are what the contributors carry forward from him. That is the lesson, and this book is our way of acknowledging it.

For Anthony, with our heartfelt thanks.

*David Orrego-Carmona, Kayoko Takeda, Ester Torres-Simón,
Nune Ayvazyan and Bei Hu.*

References

- Pym, Anthony. 1992. *Translation and Text Transfer. An Essay on the Principles of Intercultural Communication*. Frankfurt/Main, Berlin, Bern, New York, Paris, Vienna: Peter Lang. Revised edition: Tarragona: Intercultural Studies Group, 2010.
- Pym, Anthony. 1997. *Pour une éthique du traducteur*. Arras, France: Artois Presses Université; Ottawa: Presses de l'Université d'Ottawa.
- Pym, Anthony. 1998. *Method in Translation History*. Manchester: St. Jerome Publishing.

- Pym, Anthony. 2020. "Trust-based translation history: Guideline questions and an illustration". *Chronotopos – A Journal of Translation History* 2(1–2): 146–160.
- Pym, Anthony. 2027 (forthcoming). *Method in Translation History, Second Edition*. Routledge.

Prologue

Anthony Pym, on borders

ANDREW CHESTERMAN

About 15 years ago, I had the opportunity to visit Walter Benjamin's stunning memorial at Port Bou, just across the border in northern Spain, where Benjamin had ended his life in 1940, fleeing from the Nazis. I later emailed Anthony, to say how impressed I had been, and back came a message from Tarragona. Of course, he had been there, and not only been there, but checked out the details of the historical background and written an informative and personal essay about the memorial, linking his own experience of it with wider socio-political issues of exile. It was called "Benjamin at the border" and is now available via the following link: Pym, A. (2025). "Benjamin at the Border." Knowledge Commons. <https://doi.org/10.17613/nc4ge-m7k06>

Benjamin is of course well known in Translation Studies for his classic essay "Die Aufgabe des Übersetzers" ("The task of the translator"), but for Pym, embedded personally and professionally in intercultural space, Benjamin symbolized the whole concept of crossing borders – or not crossing them. Re-reading Anthony's essay now, I notice how much it reveals of his whole way of thinking and writing.

Benjamin was "at" a border when he died. Anthony writes "on" borders of all kinds, including the one at Port Bou, and has long inhabited multicultural borderlands (in Australia, Germany, France and Spain, to mention a few). The essay starts with a personal experience, changing trains at the border between France and Spain. Anthony often brings personal experiences into his academic writing: it is as if he does not quite recognize a border between the two, not an impervious one, at least. He is happy to share his experiences, for instance by sending me his essay. He contemplates the Benjamin memorial as a visitor from Australia, with a small suitcase, and compares Benjamin's experience with his own.

Other examples of this subjective involvement can be found in paratexts in his publications. His first monograph was called *Translation and Text Transfer* (1992), a work that introduces many themes that will be developed in later works, particularly in *The Moving Text* (2004). The frontispiece of the earlier work is a photo of a hand, alongside a rock painting of a second hand. The title is: "Aboriginal handprint and the author's hand; cave near 'The

Mission', Kellerberrin, Western Australia, August 1991." A similar photo appears in the 2004 volume. It makes a statement: here I stand, Anthony Pym, making personal contact with a distant time, standing in the same place as someone stood thousands of years ago, someone very different yet also similar in many ways... His hand, my hand. Connected. His second monograph, *Epistemological Problems in Translation and its Teaching* (1993) is also given a subjective frame, this time by the photo on the cover, showing two children playing on a beach. Turn the page, and you notice that they are the author's children, "busy." Read further, in the preface, and you will meet Anthony's (then) wife, whose surname takes you back to the book's publisher. And the place of publication, Calaceite, is the small village which is one of the places where Anthony has made a home. He has put down special roots in Catalonia, but not uncritical ones. Note the throw-away comment about who sponsored the memorial.

This personal way of writing shows up in other ways too: in a style that is sometimes surprisingly informal (especially in article titles: peruse his list of published articles, available on his generously informative website, and note where your eyebrows rise), and sometimes unacademically lyrical (like the very end of our essay). The informality comes partly from an underlying oral basis: some of his publications start from lectures or seminars (for instance, his book on ethics, and the one just mentioned on epistemology). Yet it is a style that can suddenly shift into abstractions: one needs to pause to think when he notes, in passing, Benjamin's "refusal of exteriority." Further down the same page, we meet the comment that the formulation of certain complex ideas "requires as much feeling as abstraction." Feeling, empathy, on one hand, and abstract conceptualization on the other: but maybe here too there is no impervious border between the two. The style is instantly recognizable. I once acted as a peer reviewer for an anonymized article, submitted to a highly reputable journal. I read the first paragraph and thought: this must be Anthony. And of course I turned out to be right.

Right from the start of our essay, there is a strong sense of history. Well, there would be, wouldn't there, from a scholar who has written several books about translation history and how to study it. We are reminded of Benjamin's personal tragedy, as a Jew, and the political background of the Nazi terror. Fittingly, the memorial at Port Bou is the work of a Jewish sculptor, Dani Karavan. (See many photos online.) But we also have more glimpses of Anthony's personal history, living "illegally" in the borderland between France and Spain, moving back and forth, with more of a sense of optimism than Benjamin had. "If you look in the right direction, there is always a way."

Anthony often seems to think in all directions at once, reaching after allusions and citations that bear witness to his wide reading. The essay includes Benjamin's description of a monoprint by Klee, *Angelus Novus*, depicting what Benjamin sees as the angel of history. (Wikipedia says that Benjamin purchased this print in 1921.) Not content with merely citing, Pym

gives us his critical reaction to Benjamin's interpretation, itself inviting a response: time is surely the primary dimension of history (one damn thing after another, as the saying goes); but the events of history do of course take place as well, in spatial dimensions.

A critical attitude is a hallmark of Anthony's work: he likes engaging in dialogue, arguing, counter-arguing, pointing out alternatives or counter-evidence, or indeed a lack of empirical evidence. The local tourist officer asks Pym if he is a philosopher, and he says no. But to what extent is this true? He is not a professional philosopher, but he has a deep interest in ideas, epistemology, and indeed moral philosophy or ethics, with two books on this topic: first in French (1997), and a revised version later in English. In terms of the philosophy of science, he is an outspoken defender of a general empirical position, grounded in evidence, logical argument, the centrality of claims and hypotheses that can be tested, and an initial attitude of fallibility – a willingness to admit that “I don't know” I may be wrong, let's look for more evidence and test further. See in particular Pym 2016. The list of publications on his homepage, by the way, notes that this article was actually published with the wrong title, which should have been “A spirited critique of a certain anti-empiricism in Translation Studies [...]”)

Being critical is easier if you have read widely and can read closely. After Klee's picture, we touch on a poem by Terry Eagleton; a painting by Eduardo Arroyo depicting the moment of Benjamin's suicide; Tzvetan Todorov, who cites 12th century Hugh of Saint Victor on borders and exiles, in a quote that Todorov (a Bulgarian in France) took from Edward Saïd (a Palestinian in the US), who had found it in Auerbach (a German in Turkey); and finally Valéry and Ovid, with a quotation found, we are told, “at the end of Lluís Marcó i Dachs, a Catalan Republican reflecting on the exile of Spanish Jews.” The range covered by this short essay is amazing, as is the wide reading and scholarship behind it, all connected, made sense of, and entertaining sense at that. And with a spot of close reading as well: it is critically suggested that Hugh of Saint Victor probably meant something a bit different from Todorov by ‘exile’: a point that is immediately linked to the EU Erasmus exchange program and the current shape of the power lines of the information age that lie along intercultural interfaces. The essay is the fruit of a mind that is curious, sensitive, reflective, seeking to understand, to make something coherent from many fragments. We cross borders fluently between aesthetics, culture, history, politics and sociology, spiraling between the personal and the more general.

Anthony came into Translation Studies with a background of work as an English teacher and a professional translator, and with a PhD in sociology. This may explain something of his interest in practical translation problems of all kinds, and above all in people, who they are and what they do, where they live, where they move, where they get their ideas, how they are treated. In some places, the sociologist notices, “you can generally talk your way around

a rule.” In other words, you can negotiate how a rule is to be interpreted. You can act as an agent, using whatever power and imagination you have. The sociologist also observes that Port Bou has largely lost its former function as a border town and must now survive as best it can, on e.g. tourism, such as people coming to see Benjamin’s memorial...

Yet an interest in people implies more than this: in Anthony’s case, it has led to many social networks-which-become-friendships, from local (in France and Spain) to European (e.g. via EST) to international (e.g. CETRA, and Anthony’s later links with Monterey and Stellenbosch, not to mention his most recent post in Melbourne).

Some of my warmest personal memories go back to the two months I spent with my wife-to-be in Tarragona, as a visiting professor at Anthony’s Intercultural Studies Group. On a free day, he and his wife took us to his book-crammed hideaway in the village of Calaceite, where he expressed horror at my query as to whether he ever felt the need to, um, throw out a book or two in order to make room for new ones... During his PhD summer school, we broke off one day to play a global game of football (Northern against Southern hemisphere), which ended in hysterical laughter. I forgot who won, but I do recall that young Marc Pym, then aged maybe 7 or 8, persuaded us all that he was to be both captain of the Southern team and also referee. (I did wonder where he got the idea from...) This was followed, or perhaps preceded, by an amazing Catalan onion party, the *calçot* feast: delicious!

As the end of the essay shows, borders can be crossed, with a good spirit.

Other references

- Pym, Anthony. 1992. *Translation and Text Transfer. An Essay on the Principles of Intercultural Communication*. Frankfurt/Main, Berlin, Bern, New York, Paris, Vienna: Peter Lang. Revised edition: Tarragona: Intercultural Studies Group, 2010.
- Pym, Anthony. 1993. *Epistemological Problems in Translation and its Teaching*. Calaceite: Caminade.
- Pym, Anthony. 1997. *Pour une éthique du traducteur*. Arras: Artois Presses Université / Ottawa: Presses de l’Université d’Ottawa
- Pym, Anthony. 1998. “Okay, so how are translation norms negotiated?” *Translation and Norms*, special issue of *Current Issues in Language and Society* 5 (1–2): 107–113.
- Pym, Anthony. 2004. *The Moving Text: Localization, Translation and Distribution*. Amsterdam & Philadelphia: John Benjamins.
- Pym, Anthony. 2016. “A spirited defense of a certain empiricism in Translation Studies (and in anything else concerning the study of cultures)”. *Translation Spaces* 5(2): 289–313.

The relevance of translation theory for translator education today

NATAŠA PAVLOVIĆ

University of Zagreb, Croatia

Amidst market-driven concerns about employability, is theory still relevant in translator education? Perceptions were sought from 62 MA students who attended, either online or on campus, a translation theory course that applied the “flipped classroom” model. Based on these insights, the author recommends including theory in translator education, ensuring a clear link to practice and fostering students’ critical engagement with both theoretical and professional issues. In this way, it is argued that theory can build future translators’ competence and confidence, preparing them for work in a changing profession.

Keywords: translation theory, translator education, flipped classroom

Introduction

This article deals with the relevance of translation theory in translator education today. Why today? What is so special about the present moment? The translation industry, together with other sectors, is undergoing a major transformation due to technological developments and the resulting changes in the workflows and business models (Sakamoto 2026). While the current technological revolution is not the first technology-related advancement to shape the translation industry (Pym and Torres-Simón 2021), the widespread application of neural machine translation and generative AI solutions is arguably having a qualitatively different impact on the translation profession, akin to that of the Industrial Revolution on manual labor (Ekpo 2025). The threat to the profession brought about by automation is compounded with accompanying trends such as commodification of translation, platformization of work, precarity, deteriorating working conditions, and a “race to the bottom” with respect to pay (Moorkens 2017; Moorkens 2020; Firat and Şahin 2026). These conditions associated with the “digital neo-Taylorist workspace” are not restricted to the translation industry, and they are not

merely new tools and business practices causing a “gradual devaluation of human experience” (Baumgarten and Bourgadel 2024: 519). Their impact on human lives is such that it is not overly dramatic to say that they invite redefining of “what it means to be human” (Ekpo 2025: 19).

What does this mean for higher education in general, and for humanities in particular? Higher education in humanities is undoubtedly in crisis, and has been for some time, struggling to attract students, improve rankings and secure funds (Guyer 2024). Some of the reasons behind the crisis in higher education are the same as those reshaping the translation industry mentioned above: the widespread use of generative AI, purporting to make even the creative aspects of academic inquiry redundant, and the neoliberal, profit-oriented policies focusing on quantifiable performance (Hayward 2024). A university degree alone does not guarantee success in the labor market, where the focus is on specific skills and adaptability. Some universities, especially those in the humanistic, liberal arts vein, tend to turn a blind eye to these trends, unwilling to “get their hands dirty” with concerns such as student “employability”², and preferring to hold onto the status quo and time-honored traditions. Such an approach leaves students unprepared for the job market and susceptible to exploitation until they acquire concrete skills in the workplace. Young people sense this and are uninterested in something for which they see no purpose. At the opposite end of the spectrum are those universities, including many translator training programs, which have undergone one reform after another in attempts to “skate to where the puck is going, not to where it is” (Welocalize 2023), with varying degrees of success. For some of them, these efforts to bridge the gap between the industry and academia have meant focusing predominantly, if not exclusively, on the practical aspects of translating, without much space in the curriculum for theory. Given what translation is today, and what a wide range of “core” and “adjacent” language services (Angelone 2026a) the industry covers, this is not surprising. With so many specific skills and competences required “to produce a transversal skill set to equip graduates for the future” (EMT 2022: 2), there does not seem to be time for “slow scholarship” (Nørgård, Bosanquet, and Grant 2024).

Innovative and creative methods such as project-based learning, simulated translation bureaus, work placements, mentorship schemes or interprofessional training have transformed translator education, bringing it closer to the real-life conditions in the industry and enabling a smoother integration of graduates in the labor market. However, theory, where it is taught, seems to have mainly retained traditional delivery methods, such as “ex cathedra” lectures. Such instructional models additionally undermine the potential usefulness of theory and research, with practitioners and students

² For a recent overview of the concept of employability in the context of translator education, see Hao (2026).

wondering how/whether theory can help them (Chesterman and Wagner 2002; Pym and Torres-Simón 2016). For students, the seemingly insurmountable divide between theory and practice makes theoretical courses detached from the realities of the professional communities they strive to join upon graduation. However, previous research has shown that students are not necessarily averse towards translation theory as such, but that their perceptions depend on how it is delivered and on the extent to which theory helps them become better translators (Pym and Torres-Simón 2016; Ordóñez López and Agost 2022).

What, then, is the place of theory in higher education today? Should it be included in study programs and how should it be taught? I will attempt to address this question by using the example of a translation theory course, which can serve as a case study for other theoretical courses in the humanities, and possibly in other fields of knowledge as well.

Methodology

The Translation Theory course that forms the background of this study is part of the Translation Stream of the graduate (MA) study program in the Department of English, University of Zagreb. This semester-long course is obligatory for the Translation Stream students (usually 20-30 per cohort), and it can also be taken as an elective course by other students interested in translation (usually no more than 5-6 students). The course starts with an overview of translation types and translation-like activities as well as a brief introduction to translation tools and resources, and basics of terminology. The content related more strictly to translation theory includes the following: introduction to translation theory and translation studies; linguistic approaches to translation – non-correspondences between language systems, equivalence theories, translation shifts, procedures and strategies, with special emphasis on cultural references, idioms, language variations (e.g. translation of dialect); functionalist approaches to translation – *skopos* theory; Descriptive Translation Studies, with emphasis on norms; cultural turn in translation studies – feminist and post-colonial approaches, and Venuti; the technological turn – the impact of technology on translation theory and practice; translation ethics. The course also features a visit by former students, who present their experiences in the translation industry since graduation, addressing student questions related to the profession and technology similar to those raised by students in Pym and Torres-Simón (2016). Outside of the Translation Theory course itself, the four-semester Translation Stream includes courses on specialized translation, administrative/legal translation, translation for the EU, audiovisual and literary translation, localization and translation technology, as well as linguistic courses and courses on anglophone literatures. Some of these courses are obligatory, but most are placed in groups from which

students choose one, according to their preferences. The study program ends with an MA thesis in the form of a research or translation project.

During the period under study, three cohorts of students attended the Translation Theory course, one in the online environment, during the COVID pandemic (winter semester of 2020/2021, 38 students) and two in the classroom (winter semesters of 2023/24 and 2024/25, a total of 56 students). Of those 94 students, 62 (66%) responded to the questionnaire survey, 30 from the online cohort (Group O) and 32 from the face-to-face cohort (Group F). No responses were excluded.

The teaching method used in both conditions was the “flipped classroom” (Bergmann and Sams 2012). I provide more details of this method and my application of it in a previous article (Pavlović 2021), where I report on the results from the online condition, or more precisely, from “emergency remote teaching” (Hodges et al. 2020), as experienced by Group O. In the current follow-up study, data related to Group F are added in order to investigate whether the same instructional model could be applied in a face-to-face environment with equal success, and what conclusions can be drawn from this. The flipped classroom model itself, while not brand new, is still novel in the higher education setting. To briefly describe the model as I used it, most lectures were displaced from the synchronous classes; instead, students were required to watch a video at home, of an average duration of around 30 minutes, and either complete a related translation assignment or participate in an online forum discussion. The synchronous classes were then dedicated to further discussion and elaboration of the topics dealt with in the videos and the assignments or online forums. The synchronous time also included additional tasks done in smaller groups (3-4 students), and then jointly commented on by the whole class. Synchronous classes took place over Zoom for Group O, and in the classroom for Group F.

The questionnaire used with Group O is provided in Appendix 2 of Pavlović (2021). Group F was administered the same questionnaire with minimal changes (e.g. the formulation “Zoom discussions” being replaced with “classroom work”). In both cases, the questionnaire was anonymous, and it was filled out by the participants at the end of the coursework and before the exam.

For the closed questions, the results were analyzed quantitatively, using JASP for the descriptive statistics and the t-tests. The responses to the open-ended questions were analyzed qualitatively, using the thematic analysis method (Braun and Clarke 2006). When cited, the participants from Group O are labelled as GO 1–30, while those from Group F are labelled as GF 1–32. Since the course was delivered in English, the questionnaire was in English, and the responses were also in English. When cited, the responses are presented verbatim, including occasional errors.

Results

Overall evaluation of the course

The overall evaluation of the course was elicited in four ways: 1) in comparison with other theoretical courses; 2) in terms of the participants' self-assessment of engagement in the course; 3) through self-assessment of how much they learned in the course; and 4) how likely they were to recommend the course to their peers.

Comparison with other theoretical courses

Firstly, the participants were asked to compare the present course with other theoretical courses they had had during their university studies. They were specifically asked to rate the delivery method used in the course rather than the course content. The participants could choose one of the following options: "the delivery of this course was much better than other theoretical courses", "the delivery of this course was better than other theoretical courses", "the delivery of this course was as good as other theoretical courses", "the delivery of other theoretical courses was better than that of this course" or "the delivery of other theoretical courses was much better than that of this course."

Fifty-two of the 62 participants rated the delivery of this course as "better" or "much better" than that of other theoretical courses, while only one person rated it as inferior to other courses (Table 1). In Group O, 28 of the 30 participants considered the course to have had a "better" or "much better" delivery than comparable courses, while in Group F, 24 of the 32 participants expressed this opinion.

Table 1: Comparison with other theoretical courses in terms of delivery

	Much better than others	Better than others	As good as others	Others better	Others much better	Total
Group O	14	14	2	0	0	30
Group F	16	8	7	1	0	32
Total	30	22	9	1	0	62

When asked to elaborate on their responses, the participants of both groups stressed the link between "theory" and "practice" that, in their opinion, this course promoted:

Most theoretical courses are purely theoretical, that is, without any tasks or assignments to make students apply the theoretical knowledge they receive in the course, which I think is not very helpful. (GO23)

Usually theoretical courses are very boring because it's mostly the professor regurgitating facts and expecting us to memorize everything,

but in this course we had example exercises, various activities, discussions, etc. so the overall experience was immensely better. (GF5)

I'm not a big fan of theoretical courses, but in this case the choice of what's important was *actually* important and pretty clearly explained. Combined with having translation tasks and discussions for certain topics, I feel like I've actually remembered a good deal of things. (GF21)

I found that I remembered things much easily and quickly, thanks to the approach, classroom discussions, relevant homework. (GF31)

Some participants explicitly comment on the instructional method as contributing to the present course being better than comparable courses:

Theoretical courses are usually pretty boring, but this one is different, because it employs other forms of teaching, not just talking about a topic. I think that group assignments and videos are the best parts of this course and are what makes better than most other theoretical courses. (GF8)

In the participants' opinion, this type of delivery makes the course more dynamic and enhances the students' motivation:

I think you did a great job with making the course more dynamic (using the methods mentioned above), so it didn't feel purely theoretical and dry, which is often the case with courses on theory. (GF20)

[In other theoretical courses] the students barely stay awake and have no motivation at all. (GF29)

A more active dynamic and interactive environment helps with staying focused. (GF32)

Unlike in the study by Ordóñez López and Agost (2022), the participants highlight the course promoting critical thinking, as illustrated by these comments:

Other theoretical courses usually consist of memorizing what other people said in a certain period of time on a certain topic, while this course focused on critical thinking. (GO15)

The goal of the course was not to memorize the definitions, classifications or authors, but to think critically about everything. [...] I think that because of such a model of classes I will retain everything we did for much longer in my memory than if I'd only had to cram. (GO26)

Self-assessment of student engagement

Table 2 shows the participants' self-assessment of engagement in the course. In total, 39 of the 62 participants were actively engaged, either often or occasionally, while the remaining 23 listened but rarely spoke. One of the 19 Group F participants who chose the second option ("Listened actively and occasionally spoke"), added a note explaining that the description was mostly true of them, but that there were also days when their mind was completely elsewhere. In the interest of transparency, this is indicated in Table 2, in brackets.

Table 2: Participants' self-assessment of engagement

	Often actively participated	Listened actively, occasionally spoke	Listened but rarely spoke	Listened only "with one ear"	Did not listen
Gr. O	4	13	13	0	0
Gr. F	3	19 (18)	10	0	0 (1)
<i>Total</i>	7	32	23	0	0

Self-assessment of learning

The participants were asked to assess how much they had learned about translation in this course on a 1–5 scale (1 = "almost nothing new or useful"; 5 = "a lot of new and useful stuff"). As can be seen from Table 3, only two participants chose the middle point, and no one gravitated towards the lower end of the scale. The differences between the groups were minimal.

Table 3: Self-assessment of learning

	Point 1	Point 2	Point 3	Point 4	Point 5	Total
Group O	0	0	1	7	22	30
Group F	0	0	1	10	21	32
<i>Total</i>	0	0	2	17	43	62

Recommendation of the course

The fourth element of the participants' overall evaluation of the course was obtained when they were asked to what extent they would recommend the course to colleagues who wished to become translators. Sixty-one of the 62 participants strongly agreed or agreed that they would recommend the course even if it were not obligatory, with Group O agreeing slightly more strongly (Table 4).

Table 4: Recommendation of the course

	Point 1	Point 2	Point 3	Point 4	Point 5	Total
Group O	0	0	0	4	26	30
Group F	0	0	1	8	23	32
<i>Total</i>	0	0	1	12	49	62

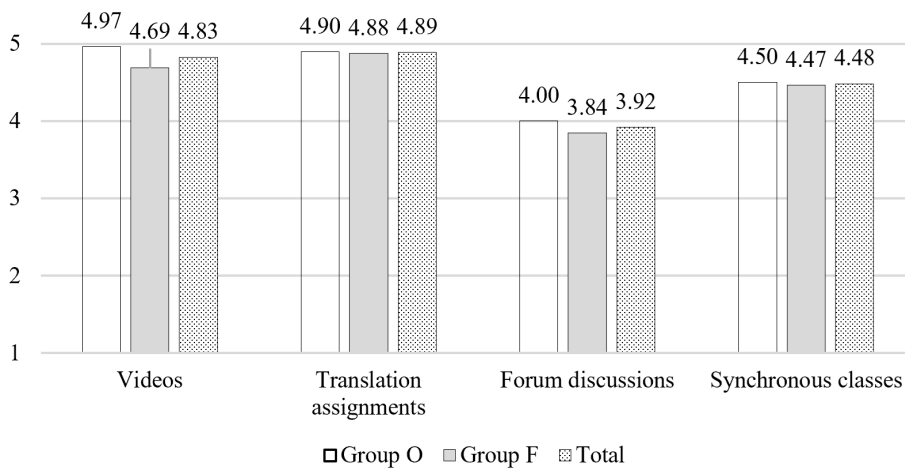
Evaluation of course components

Quantitative evaluation of course components

One of the questions in the survey asked the participants to rate the four components of the course described in Section 2: the videos, the translation assignments, the forum discussions, and the synchronous classes via Zoom (Group O) or in the classroom (Group F). The usefulness of each element was rated on a Likert-type 1–5 scale, where 1 meant “not at all useful” and 5 “very useful” (Figure 1).

Figure 1 shows the results for Group O (white), Group F (grey) and the combined results (pattern). As can be seen, the videos and the translation assignments received the highest scores (mean = 4.83 and 4.89 respectively), while the forum discussions were rated slightly lower (mean = 3.92). We look at the reasons for the latter result below. The synchronous classes were also rated very positively, as useful to very useful (mean = 4.48).

Figure 1: Evaluation of course components



The chart also shows that participants in Group F rated all the course components slightly less enthusiastically than their counterparts in Group O. In the case of the videos, the difference was statistically significant on an independent two-sample t-test ($p < .05$). Group O attended the course in extraordinary circumstances, with both their studies and their lives in general being severely constrained by the pandemic. We can therefore speculate that the use of an innovative teaching method would have meant more to them than to their colleagues studying under normalized conditions. In this respect, the results of Group F can be thought of as more realistic. It is indicative that an important aspect of Group O’s impressions of this course was related to the perceived “support” they received in the course (Pavlović 2021), while this

notion was not brought up, either directly or indirectly, in the responses of Group F.

In any case, the overall mean scores ranging from 3.92 for the forum discussions to 4.83 for the videos can be described as a very positive evaluation. Below, I present the qualitative results related to each of the components.

Qualitative evaluation of course components

The videos. Among the advantages of the videos, participants highlighted their availability, and the fact that they could be re-watched at any time (GO4, GO16, GO17, GO30, GF8, GF10). The videos were described as entertaining (GO4, GO9, GO19, GO28, GF13), informative (GO7, GO29), creative (GO18, GF29), shorter than traditional lectures (GO29), and easier to follow (GF13). Here are some statements about the videos in the participants' own words:

I liked being able to watch the video lectures (and the fact that I could watch them multiple times) (GO17)

I found the videos to be a great way of presenting course topics. I think they definitely made the class more fun and they are easier to follow than standard lectures. (GF13)

The videos are extremely helpful because we can watch them on our own time and thus understand the matter better. (GF10)

[The videos] were very creative and I was excited to see what you would come up with next! (F29)

The participants spontaneously brought up the role of the videos in freeing up the time for discussion in contact lessons, which is the main tenet of the flipped classroom model:

The videos were a great way of covering the theoretical part of the course, while leaving time to cover concrete, practical tasks and examples in the classroom. I believe it is more useful to use the in-person lessons for discussions. (GF11)

The videos were very useful because I could watch them when most convenient, and they allowed us to be prepared for class so we could just talk about the material. (GF5)

I really like that the professor had videos we had to watch at home and in class we only discussed real life situations that were relevant to that video rather than only studying theory. (GF19)

Only one participant, from Group F, said they would have preferred to be given reading assignments to videos.³ As described in Pavlović (2021), the videos used in this course were made specifically for the course by the author, in the COVID-19 pandemic conditions. In the absence of such content, any available videos could be used, or indeed reading material. Since previous research has shown that students may find it a challenge to benefit from translation theory when the “density and volume” of reading materials is perceived to be too high (Ordóñez López and Agost 2022), the videos might be considered a suitable medium for facilitating and motivating students’ engagement with complex theoretical concepts. This approach can be used to nudge students towards tackling more dense research articles when they have built appropriate foundations.

The translation assignments. The participants place the translation assignments among the most useful and enjoyable activities done in the course, stating that they helped them to understand and retain theoretical knowledge, as well as to apply it to real-life situations:

The assignments usually nicely complemented the videos and helped me see how the knowledge provided worked in practice. (GO19)

Assignments involving translation tasks are great and a useful way of showing how theory can be used in actual translations. (GF4)

Assignments were also very useful because it was easier to understand the theory after applying it to a particular task. (GO25)

I think this class has a very good balance of the more straightforward theory learning and practical assignments which help put that theory to use. (GF12)

It is much more useful to make the theory just the basis for doing practical tasks rather than theory being the only thing discussed and worked on in the course. (GF11)

The assignments also served to develop critical thinking, problem-solving skills, and creativity:

Practical assignments encourage problem solving and critical thinking and they can help us with recognizing the learned material in real-life situations. (GO4)

[The assignments] let us show our creativity and were useful in giving insight into different translation techniques. (GF15)

The coursework was very diverse and encouraged creativity. (GF6)

³ It should be mentioned that, for one of the topics (cultural turn in translation studies), a reading assignment was used instead of a video.

Most of these assignments involved experimental translations and translations that took the students outside of their comfort zone, such as song translation, translation of dialect, translation of idioms involving word play, translation of advertisements, translation of a text with many cultural references, etc. This was done in a safe environment and without the pressure of grades, since the assignments were obligatory for the students, but not graded as such. In other words, “the students translate as a process of testing and discovery, not as a training practice in itself” (Pym and Torres-Simón 2016: 199). Such experiments with unpredictable outcomes are something students will be rarely able to afford once they have graduated and started working for clients, but they can nevertheless be seen as an “upfront investment that can ultimately optimize work performance” (Angelone 2026b: 29).

The forum discussions. Forum discussions were used in this course to accompany the videos, especially those dealing with content less amenable to application through translation assignments (e.g. norm-based theories, cultural turn in translation studies, translation ethics, the impact of technology). The students were asked to watch a video and respond to a set of questions posed by the teacher to prompt discussion or ask their own questions and respond to their colleagues’. The main issues were then taken up in the synchronous classes.

The forum discussions provided the opportunity for students to hear the opinions of their peers:

The forum assignments were really interesting, because I was able to see my colleagues’ opinions and comments on various subjects and learn a lot from them as well. (GO10)

The forum discussions and creative translation tasks were really interesting and fun to do. (GF31)

I usually don’t like these types of assignments, but the ones done in this class were very good [...]. It was also interesting to read about different opinions and experiences my colleagues had. (GO30)

I liked the forum discussions; I thought they would be stilted and a waste of time, but I ended up really enjoying reading my colleagues’ ideas and responding to them. (GO23)

They were also a suitable venue where those who do not like to speak in front of the group could express their views:

[Forum discussions] really motivated students like me who struggle with class participation to still express their thoughts and opinions. (GF6)

At the same time, the participants' responses to the open-ended questions help explain why the forum discussions received slightly lower marks than the other three types of activities used in the course:

Forum assignments were very interesting because they made me think more deeply about the course material, but sometimes it was a bit difficult to follow the discussion, especially when a lot of answers were posted at the same time (at the last minute). (GO25)

I was always overwhelmed by the number of opinions and responses, that I never managed to give them a proper look. (GO14)

The forum assignments were doomed to be repetitive by nature. Our opinions didn't vary that much, everyone kind of copy-pasted responses from other students and re-did them a bit. (GO28)

As far as forums go, I feel like they were a bit less useful for me personally, because I often did the assignments in a rush and wouldn't take time to read everyone's input. (F8)

I dislike the forum option because the discussions become too official, more focused on the way things are said and less on what is actually being said. (GF18)

[I]t did sometimes feel a bit unnecessary to take part in the forum discussions since other students had already mentioned many points and I couldn't find anything new or interesting to add. (GF20)

One of the problems associated with the forum discussions was the platform on which they were conducted. Students seem to be used to more modern-looking user interfaces than those offered by the Moodle e-course, as can be seen from these comments:

Due to the [Moodle] forum being somewhat clumsy, the discussions in the forum came off rigid at times, with people talking past each other and repeating some of the points that had already been made. (GO12)

Upgrades to the use of the forum would be beneficial, but that has more to do with [Moodle] as a platform. (GO1)

Forum discussions are a good idea, but for me personally, [Moodle] forums are not quite ideal for a developed conversation. (GF10)

Alternative platforms should be sought to facilitate written discussion among course participants, as it has the potential to enrich the exchange of knowledge and prepare the ground for oral discussion in the synchronous classes.

The synchronous classes. Regarding the synchronous component of the course, the comments were overwhelmingly positive, with the participants being pleased with the "relaxed" (GO10, GF6), "positive" (GO10), "nice and friendly" (GO10), and "welcoming" (GO19) atmosphere, which fostered

interaction and engagement (GF15). The participants furthermore welcomed the opportunity to hear different perspectives and opinions from their peers. These are some of the comments about the synchronous classes:

Classroom work and discussions were also very productive, and the relaxed atmosphere made them more enjoyable. (GF6)

The positive atmosphere was probably the biggest plus in my book, because it allowed me to be relaxed and attend lectures without a knot in my stomach (which is not the case with other subjects). (GO10)

The atmosphere was always nice and friendly and it felt like we were all hanging out while talking about translation theory. (GO10)

I like that there was more interaction in this course. By discussing the topic we were able to really think about the information we got and it was more engaging. (GF15)

Although I did not always have something to contribute to the discussions, I found them very useful because I heard many new perspectives on translation issues that I had not considered before. (GO12)

[Y]ou made Zoom a very welcoming space and gave everyone a chance to speak. I felt heard and like my opinions and thoughts mattered and it was nice to hear colleagues express their own views and opinions. (GO19)

There was also a lot of feedback and in-class discussion which I think helped me fix certain things I was doing wrong and see different perspectives. (GF22)

I liked the discussions via the forum and Zoom because, when I had to think about a topic and develop my own opinion, it became clearer straight away, and generally in this way I was able to remember the content better, when I had to think about it. (GO26)

As reported in Pavlović (2021), some participants who attended synchronous classes in the online environment mentioned that the quieter students were not able “to get a word in” (GO20) because the more assertive students tended to dominate the discussion. It was suggested that the teacher should moderate the discussion more actively (GO4). On the other hand, calling out students by name to participate in discussions can be a source of anxiety for some students in both the online and classroom environments, as witnessed by these two comments respectively:

What I loved most about [the Zoom discussions] was that the professor did not call us out by name to answer questions, we could participate in the discussions freely, how, when, and if we wanted (this was actually my favorite Zoom class because of that very reason, I felt very anxious and on edge in the rest of the zoom lectures in other subjects, and that

experience influenced my online classes experience the most, in a negative way). (GO10)

Thank you for not calling out names during in-class discussions (GF18)

It can therefore be a challenge for the teachers to strike the right balance between being too intrusive or too laissez-faire in their approach to discussion moderation, with personal preferences and character traits playing an important role.

The perceived benefits of the course

Based on the responses to open-ended questions, the major benefits that the course offered can be divided into two overarching categories: the immediate usefulness of theoretical knowledge for the improvement of translation competence, and the forward-looking relevance for the future.

The first category encompasses the acquisition of the conceptual apparatus necessary for thinking about and discussing translation (putting thoughts into words, learning how to back translation solutions with arguments); broadening one's perspective and viewing translation in a new way; critical thinking; the ability to apply theories in practice; and confidence in applying theory. Here are some illustrative examples:

[The course] has been useful by helping me put the ways I translate into words and opening my eyes to other ways something can be translated. (GF15)

I've learned a lot about the theoretical background of translation [...], which helped me learn how to express my own opinion with proper arguments and examples. (GO14)

It was interesting to learn about translation theory and the various ways scholars have described the process of translation. It challenged a lot of preconceived notions I had, and I think I now have a much better picture of what translation is. (GO25)

[The course] made me realise that there is much more to translation than just translating. (GO15)

It made me start thinking about translation in different ways and having more of an open mind. (GF5)

It reminds the student that they are part of a scientific community and invites them to approach the theory in a critical way (GO12)

It made us think about how we can apply in practice all the theoretical knowledge that we gained. (GO30)

I was introduced to actual procedures and strategies used by translators (which I have used before in my studies, but wasn't aware of them)

that made me feel more competent and equipped as a translator. (GF31)

It broadened my scope of the translation profession. It helped me to be more professional and thoughtful, and enabled me to explain my approach to translation. (GF1)

The second category of benefits from the course relates more to the future and the relevance of the course for the real world. This category includes points such as being prepared for the future (particularly how to handle challenging situations, guidance in choosing the future area of specialization); expectations of the translator's role and responsibilities; and insights into the development of the profession, especially as related to the impact of technology. The following comments are illustrative of this category:

I learned a lot about Translation Studies in general and the theory behind my future profession. I believe this will help me choose what line of work I wish to do once I graduate. (GO11)

I gained a clearer picture of what I would like to do once I graduate. (GO30)

I liked that the tasks were relevant to real life situations in which we will find ourselves. I really feel this class prepared me for the future, in some aspects, even better than the translation exercises. (GF19)

[The course] served as an excellent introduction to translation theory, it introduced key concepts and issues that I will continue to use and think about throughout my translation experience. It also made me think about the role and the accountability of the translator. (GO12)

I became aware of what is expected [...] and how to deal with complicated situations we can experience as translators. (GF19)

I learned about some useful concepts that I find will be helpful in my future endeavours. (GF17)

It was useful in that a lot of practical examples were offered, which helped me to get acquainted with the real-world situations in which I may find myself one day. (GF23)

I thought a lot about translation and language development and technological development, and how it will influence our career in the future. (GF7)

It is clear from these comments that the course has helped students gain a better knowledge of translation as their chosen profession and to improve their confidence as translators, which are crucial factors in the development of their professional identity (Singer 2022).

Limitations of the study

The present study, while offering rich data on student perceptions of the course under investigation, also has many limitations. One of them is the self-selection bias: while the response rate of around 66% is quite good, it is possible that students who were more satisfied with the course were more inclined to complete the questionnaire, leading to more favorable results. Furthermore, like other self-report studies, this study relies solely on the participants' perception of the course benefits for their translation competence, the development of which was not assessed for the purposes of this study. Finally, this was a case study involving a single course with a single instructor in a very specific setting and no attempts at controlling the variables.

Conclusion

The results of this study have shown that translation theory can be highly appreciated by students, who can see its usefulness and its relevance at a time when they are justifiably concerned about employability and the future of the profession. Even the most informed teachers and curriculum developers cannot predict the technologies that will be used in the future, the skills they will require and the forms of employment they will beget. But, to cite Angelone (2026b: 28), one thing that can be expected to remain constant amidst incessant change is “the need for translators to critically reflect on, assess, and optimize their own performance.” Or, in the words of Pym and Torres-Simón (2016: 197), “Sooner or later, some creativity and critical thought is called for. And stimulations for inquiring minds, from research or theory, can nudge those processes along.” I would therefore argue that, regardless of what translation profession will look like in the future – or indeed, what other career path students will end up choosing – having the ability to tackle complex theoretical concepts and apply them in practice, to think critically, back their solutions with arguments, and make ethical decisions will help students navigate the challenges of the changing labor markets. These are precisely the advantages of a humanities university education that are coming to the fore today, in the era of overly narrow specializations and uncritical use of technological solutions in workplaces.

My experience with this course has shown that the divide between theory and practice is largely a construct bred from years of suboptimally designed curricula and not an inevitable dichotomy. Rather than opposites, theory and practice can be mobilized as complementary types of engagement with knowledge creation, with more theoretical or more practical aspects being brought to the foreground at different points within the course or indeed a single class or assignment. Such an approach builds on Schön's (1987) notion

of “reflective practice”, which has been very influential in my own development as a teacher. As Pym and Torres-Simón (2016: 199) say, “theory is no longer a set of irrelevant things to learn but becomes, in the more optimistic scenario, a way of enabling discussions [and] reasoned dissent.” However, these two authors argue for a “limited role of theory” (in the sense of “theorization”), which is marshalled at points when students run into difficulties or conflicting perspectives, “in a process that need be pursued only for as long as the learning collectivity has interest and energy” (Pym and Torres-Simón 2016: 199). I agree that such an approach is indeed resourceful, but it might be better suited for courses at lower levels (e.g. BA) or in situations where the course has a very limited number of contact hours. I would argue that the present study has shown that we can expect more from students at the MA level, especially when there is enough time in the course to tackle whole theories rather than only some theoretical concepts that may be useful for the task at hand.

However, what this study has also shown is that students’ positive views of translation theory depend to a large extent on the instructional model used in a course. In this paper, I have argued that a model which fosters student engagement, autonomy and experimentation, blurring the lines between theory and practice, such as the flipped classroom model described here, is closely related to students’ acceptance of a theoretical course. The flipped classroom model has the advantage of being easy to adapt to the course content, the number of hours, the students’ needs, the teacher’s preferences, and so on, but this is not to say that other innovative pedagogical designs could not be used in its place. With universities facing a crisis, both curriculum developers and the teaching staff will have to devise creative ways to keep higher education relevant in the years to come.

Declaration of AI use

The author did not use any AI-assisted tools in the preparation of this manuscript.

References

- Angelone, Erik. 2026a. “Diversification and portfolio careers”. In C. Walker and J. Lambert (eds) *The Routledge Handbook of the Translation Industry*. London and New York: Routledge. 307-320.
- Angelone, Erik. 2026b. “Generative AI as a facilitator of deliberate practice in translator training”. In J. C. Penet, J. Moorkens and M. Yamada (eds). *Teaching Translation in the Age of Generative AI: New Paradigm, New Learning?* Berlin: Language Science Press. 27-47.

- Baumgarten, Stefan, and Carole Bourgadel. 2024. "Digitalisation, neo-Taylorism and translation in the 2020s". *Perspectives* 32(3): 508-523.
- Bergmann, Jonathan, and Aaron Sams. 2012. *Flip your classroom. Reach every student in every class every day*. Arlington, VA: International Society for Technology in Education.
- Braun, Virginia and Victoria Clarke. 2006. "Using thematic analysis in psychology". *Qualitative Research in Psychology* 3(2): 77-101.
- Chesterman, Andrew, and Emma Wagner. 2002. *Can Theory Help Translators? A Dialogue Between the Ivory Tower and the Wordface*. London and New York: Routledge.
- Ekpo, Eno Ubong. 2025. "The role of philosophy in translators' response to automation". *Bridge: Trends and Traditions in Translation and Interpreting Studies* 6(1): 8-28.
- European Master's in Translation (EMT). 2022. *Competence Framework*. https://commission.europa.eu/system/files/2022-11/emt_competence_fw_2022_en.pdf.
- Firat, Gökhan, and Mehmet Şahin. 2026. "Platform economy models in the translation industry". In C. Walker and J. Lambert (eds) *The Routledge Handbook of the Translation Industry*. London and New York: Routledge. 290-306.
- Guyer, Sara. 2024. "Director's Report: Strategies for Persistence in a World under Duress". World Humanities Report, Consortium of Humanities Centers and Institutes (CHCI). <https://worldhumanitiesreport.org/wp-content/uploads/2024/10/WHR-DirectorsReport.pdf>
- Hao, Yu. 2026. "Employability and translator education". In C. Walker and J. Lambert (eds) *The Routledge Handbook of the Translation Industry*. London and New York: Routledge. 496-510.
- Hayward, Katy. 2024. "Machine unlearning: AI, neoliberalism and universities in crisis". *Red Pepper* 245. <https://www.redpepper.org.uk/culture-media/technology/machine-unlearning-ai-neoliberalism-and-universities-in-crisis/>
- Hodges, Charles, Stephanie Moore, Barb Lockee, Torrey Trust, and Aaron Bond. 2020. "The difference between emergency remote teaching and online learning". *EDUCAUSE review*. <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remoteteaching-and-online-learning>.
- Moorkens, Joss. 2017. "Under pressure: translation in times of austerity". *Perspectives* 25(3): 464-477.
- Moorkens, Joss. 2020. "'A tiny cog in a large machine'. Digital Taylorism in the translation industry". *Translation Spaces* 9(1): 12-34.
- Nørgård, Rikke Toft, Agnes Bosanquet, and Barbara M. Grant. 2024. "The slow scholar in the accelerated university: Slowness as solidarity". In P. Gibbs, V. de Rijke, and A. Peterson (eds) *The Contemporary Scholar in Higher Education*. London: Palgrave Macmillan. 133-153.

- Ordóñez López, Pilar, and Rosa Agost. 2022. "Future translators' views on translation theory: a qualitative approach". *The Interpreter and Translator Trainer* 16(2): 158-176.
- Pavlović, Nataša. 2021. "'It felt like we were all hanging out while talking about translation theory': Lessons learned from a flipped translation theory course in emergency remote teaching during the COVID-19 pandemic". *Current Trends in Translation Teaching and Learning E* 8: 31-85.
- Pym, Anthony, and Ester Torres-Simón. 2021. "Is automation changing the translation profession?" *International Journal of the Sociology of Language* 270: 39-57.
- Pym, Anthony, and Ester Torres-Simón. 2016. "Designing a course in Translation Studies to respond to students' questions". *The Interpreter and Translator Trainer* 10(2): 183-203.
- Sakamoto, Akiko. 2026. "Changing practices in the translation industry". In C. Walker and J. Lambert (eds) *The Routledge Handbook of the Translation Industry*. London and New York: Routledge. 26-39.
- Schön, Donald. 1987. *Educating the Reflective Practitioner*. San Francisco: Jossey-Bass.
- Singer, Néstor. 2022. "How committed are you to becoming a translator? Defining translator identity statuses". *The Interpreter and Translator Trainer* 16(2): 141-157.
- Welocalize. 2023. "Embracing disruption in the language services industry". *Slator*. <https://slator.com/embracing-disruption-in-the-language-services-industry/> Visited May 2026.

Tracing the evolution of translator training in Anthony Pym's research

VOLGA YILMAZ-GÜMÜŞ

Anadolu University, Türkiye

Anthony Pym, as a dedicated translator trainer and translation researcher, has produced a substantial body of work that collectively sheds light on the major developments, debates and turning points in translator education over the last three and a half decades. This study sets out to provide a meta-look at Pym's contributions to translator-training research, based on a systematic reading of over fifty works published between the early 1990s and the present. Tracing the evolution of translator training through Pym's work, this meta-look also seeks to show the relevance of his ideas for current practices and future directions in the field.

Keywords: Anthony Pym, translator training, translation competence, translation and technology, translation profession

Introduction

As a researcher always interested in the evolving and complex nature of translator training (TT), I have found Anthony Pym's ideas on translator education inspiring and responsive to the challenges in the field. He is one of the most influential and prolific researchers in TT, having authored or co-authored, to the best of my knowledge, 54 publications on the topic. These publications collectively help us trace the major developments, debates and turning points in TT from the early 1990s to the present. This study sets out to provide a meta-look at Pym's contributions to TT research, documenting how TT has evolved in higher-education contexts worldwide and bringing into focus the main debates in TT from his perspective. The analysis is based on a systematic reading of his publications on TT, including books, peer-reviewed articles and book chapters. Rather than aiming for an exhaustive bibliographical review, the study focuses on recurring themes and concerns throughout Pym's works, especially distance learning, competence, technology and the changing sociological aspects of translator education. The

themes discussed in the analysis were determined through close reading, comparison and manual clustering of related topics in the corpus.

A thematic look at Anthony Pym's works on translator training

Before providing a detailed thematic analysis, it is useful to present a quantitative overview of the main topics addressed in Pym's works on TT. It should be noted that this study focuses on works closely related to the training of translators. However, many of Pym's other publications also offer important implications for TT, even when this is not their main focus. Among the publications examined in this study, technology-related themes constitute the largest group (15 works), reflecting his interest in translation technologies since the late 1990s and, more recently, machine translation (MT) and generative artificial intelligence (GenAI). This is followed by studies focusing on employment, professionalization, and market-related issues (9 works), also with a strong focus on the changing effects of technology. Translation competence and distance/e-learning each constitute four works. Other TT-related studies cover a wide range of areas, including translator trainer training (3 works) and the relationship between translation and language teaching (4 works). Although these categories inevitably overlap in some, if not most, cases, the distribution provides a general sense of how developments and debates in TT are represented in his work.

Distance learning and e-learning in translator training

I first met Anthony Pym in 2006 when I began the PhD Program in Translation and Intercultural Studies at the Universitat Rovira i Virgili (URV), which was offered remotely with regular gatherings of students and professors in Tarragona. I was very happy to be accepted to the program; however, many people around me in Turkey were skeptical about my decision to enroll. At the time, distance education was widely viewed with suspicion and often dismissed as lacking academic quality. Nevertheless, the program at the URV not only offered a rigorous academic environment but also brought together students and professors from different parts of the world, creating an intellectually stimulating community that was anything but *distant*.

The doctoral program in Translation and Intercultural Studies started in 2003. In the early 2000s, Pym authored and co-authored several papers on distance learning and e-learning (e.g., Pym 2000a; Pym 2000b; Fallada-Pouget and Pym 2000; Pym 2002a). Anthony Pym views the "distance" in distance learning not as a deficit, but as a powerful pedagogical asset. He identifies three key types of distance: external (physical separation), internal (hierarchical differences) and cultural (the gap between the learning activity and a student's everyday life) (Pym 2000a). Pym (2000a) argues that

maintaining an optimal “Vygotskian threshold” of distance provides students with enough novelty to stay motivated, while remaining close enough to ensure comprehension. Breaking from daily routines through this distance is often precisely what makes education inspiring and transformative.

A main principle of Pym’s approach is the recognition that online learners differ significantly from traditional on-campus students, as they are often older, working professionals managing complex schedules. Because of this demographic shift, Fallada-Pouget and Pym (2000) argue that directly comparing distance education with face-to-face learning to find a “no significant difference” is a fundamentally flawed research methodology. Instead, qualitative attention must be paid to the unique challenges these learners face, particularly student distress, which frequently derives from technological frustration, ambiguous expectations and a sense of isolation.

That being said, to combat this distress and prevent waning motivation, Pym (2002a) advocates interactive, community-driven online environments built on the concept of “controlled asynchrony.” Rather than letting students work entirely at their own isolated speeds, this approach keeps cohorts on the same weekly schedule to facilitate meaningful peer collaboration and socialization. Pym (2000a, 2002a) also strongly advocates tandem learning, in which students with complementary linguistic or technical skills rely on one another to complete tasks.

Over two decades later, with the global COVID-19 pandemic, distance and online education have become an indispensable part of contemporary life. And it seems they are here to stay, increasingly complementing and, in some cases, replacing traditional face-to-face contexts. Today, we can feel the implications of Pym’s approach in modern distance education. Rather than asking if online learning is better or worse than in-person formats, today’s educators focus on matching technological tools with specific instructional goals that cater to the needs of online learners. Probably, Pym’s most important lesson is that all education relies on human-to-human communication; digital tools must be used to expand, rather than restrict, the human aspect of the learning experience.

Towards transferable skills in translator training

Translation competence has been one of the most popular and controversial topics in translator-training research. Several typologies of competence have been developed, some of which are still used and regularly updated in light of the profession’s evolving nature and pedagogical tendencies. Pym (1991) offers an early critique of multicomponent approaches to competence, an argument that reappears in his later works:

[T]he multicomponent models of competence are heavy with assumptions not just about what translation is and how it should be taught, but more especially about the level at which specific teaching is

needed, and for how many years. They inevitably feed into complex professional profiles (“a good translator needs A, B and C...”); they thus underscore not just a transcendental ideal translator who has no place in the fragmented market, but also the long-duration interdisciplinary training programs that purport to produce such things (mostly university degree programs lasting four or five years). In most cases, the complex models of competence coincide more or less with the things taught in the institutions where the theorists work. (Pym 2003: 487)

He argues that complex models, which often list exhaustive linguistic, cultural and technological sub-competencies, tend to reflect institutional curriculum structures rather than the realities of a fragmented professional market. Defining competence in an electronic age, Pym (2003) distills translation into a two-fold process based on decision-making and problem-solving: the ability to generate multiple viable target texts for a single source text, followed by the ability to select the most appropriate one:

We propose that, together, these two skills form a specifically translational competence; their union concerns translation and nothing but translation. There can be no doubt that translators need to know a fair amount of grammar, rhetoric, terminology, computer skills, Internet savvy, world knowledge, teamwork cooperation, strategies for getting paid correctly, and the rest, but the specifically translational part of their practice is strictly neither linguistic nor solely commercial. It is a process of generation and selection, a problem-solving process that often occurs with apparent automatism. (Pym 2003: 489)

This definition has gained even greater relevance in an age that has moved beyond being purely electronic and is increasingly dominated by artificial intelligence (AI). As ongoing debates question whether human translators are or will be replaced by AI, translation competence can be understood as the ability to select one viable target text from amongst the countless options generated by AI tools and to adapt it for the target audience. This, however, raises a basic question: can one meaningfully post-edit AI output without knowing how to translate? More than two decades later, it therefore seems more important to keep in mind that translation competence first and foremost concerns “nothing but translation”:

Most of the electronic tools are simply techniques that speed up and broaden the production of alternative TTs (cf. what can be done with web searches); others are extensions that favour the authoritative elimination of alternatives (cf. the basic advantages of translation memories and glossaries of all kinds). Translators produce and select from alternatives; the various new technologies do not alter those tasks, they simply make them wider-ranging, dealing with more of the world in less time. To that extent, a minimalist view of competence

should help keep us aware of the ends of our tasks, without getting lost in the means. (Pym 2003: 489)

Ten years after defining translation competence in an electronic age, it should not be a coincidence that, in the *Meta* journal again, he sets out to define translation skill sets in a machine-translation age. Neural machine translation (NMT) had not been introduced yet, and Pym suggests revising our translation competence models and training programs in consideration of MT:

I would argue, statistical-based MT, along with its many hybrids, is destined to turn most translators into posteditors one day, perhaps soon. And as that happens, as it is happening now, we will have to rethink, yet again, the basic configuration of our training programs. That is, we will have to revise our models of what some call translation competence. (Pym 2013: 488)

The technological changes related to translation have always required changes in various aspects, including the basic terms of translation, e.g., the source text renamed as the start text given that translators no longer work with a single original text but with “the text, the translation memory, the glossary, and the MT feed” (Pym 2013: 492), and the social space of translation in which machine translation, non-professional post-editors and professional translators collaborate, requiring new hybrid skill sets beyond traditional models of translation competence (Pym 2013: 492).

In the age of MT, translators need skill sets that include the ability to quickly master new technologies, assess the reliability and cost-effectiveness of proposed matches, allocate cognitive effort according to task conditions and pay structures, detect suprasentential errors such as cohesion and punctuation problems caused by segmentation and conduct substantial post-draft revision (Pym 2013: 494-497). Considering that technologies have been changing rapidly, teaching MT should be conceived not as a standalone technical module but as a skills-based pedagogical approach integrated across the curriculum, emphasizing repeated hands-on use of technologies, peer collaboration, self-analysis of translation processes, teamwork with area experts and the development of adaptive expertise rather than static tool-specific knowledge (Pym 2013: 497-499).

These were the studies that directly address competence in Pym’s research. Pym (2015) also defines translation in terms of risk management, i.e., as a process affected by three forms of risk, including credibility risk, uncertainty risk and communicative risk, which emphasize that translators constantly make decisions under imperfect conditions. This perspective is still in line with his earlier definition of competence as the ability to select a viable translation solution “quickly and with justified confidence” (Pym 2003), since such confidence necessarily involves evaluating uncertainty and assuming responsibility for one’s choices. Furthermore, works such as Pym (2021), Pym et al. (2023) and Macreadie et al. (2025) discuss trust, primarily in

relation to credibility risk, as an essential element of translation practice. These works suggest a broader understanding of competence that extends beyond decision-making only related to the text: "...confronted by alternatives between which an ethical decision is to be made, translators should look at least beyond the text in front of them" (Pym 2021: 21).

Taken together, Anthony Pym's ideas on competence in TT have focused first on moving away from multicomponent models, which he considers insufficiently adaptable to different training contexts and to a fragmented market structure. Second, as technological developments are redefined and continue to redefine what is expected of translators, he places increasing emphasis on transferable skills that enable translators to adapt to new tools, workflows and roles, as well as on ethical decisions that require them to build trust with others and to manage and take risks.

Technology without technological determinism

As seen in the above works on competence and skills, technology has been a consistent theme in Anthony Pym's work. Rather than seeing technology as a disruption, they emphasize the processes of change, conditional effects and pedagogical responses. In his works, technology is regarded not as an autonomous force, but as a socially embedded factor whose impact depends on how it is integrated into professional practice and education.

Pym's collection of works indicates that his interest in technology began in the early 1990s, with discussions on the role of technological skills in defining competence. In line with technological developments in translation, two major clusters of works stand out in terms of both period and topic. In the first decade of the millennium, when translation technologies began to be integrated into TT, he authored and edited several works directly related to their teaching (Pym 2005, 2006, 2008a, 2008b, 2011). The second cluster, comprising works published from the second decade of the 2000s to the present, focuses primarily on MT and GenAI, particularly their implications for TT and the translation profession.

What Pym (2005) wrote over two decades ago remains highly relevant today, as the ethical aspects of MT and AI in translation are at the very heart of debates in both academia and the profession: "Train people to work not just with technology, but against it as well." We will come back to this later.

Pym (2006) drew attention to the additional asymmetry created by the inclusion of technologies in the translation classroom and the transfer of authority from teachers to technology, especially computer-assisted translation (CAT) tools, relating this to the already asymmetric relationship between teachers and students. In this paper, Pym proposes practical spatial solutions, arguing that "empowerment begins in architecture" (2006: 120), as well as time-related strategies to foster three-way interaction among teachers, students and technologies. However, within this sphere of interaction, he

emphasizes that it is equally important to recognize when technologies should be set aside. In particular, he emphasizes the need to “select only one viable” target text from the expanding range of alternatives generated with technological tools, returning to his discussion of competence (Pym 2003).

Pym (2008a, 2008b) provides a wide range of practical solutions for preparing students for the profession, listed under cultivating risk awareness, teaching the effective use of technology and diversifying professional roles. First, educators must develop students’ awareness of communicative risk distribution so they understand that not all parts of a text are equally important; this can be achieved through practical exercises such as text summarization, spoken face-to-face translation and peer product-testing. Second, rather than dedicating excessive class time to specific, rapidly changing software, institutions should teach students how to independently learn and evaluate new tools, using web-based resources and real-world work placements to build technological adaptability. Finally, because pure translation is no longer the sole focus of the industry, programs should train students for a variety of roles by teaching controlled writing, revision strategies, corpus and project management and collaborative teamwork utilizing shared translation memories. To successfully implement these comprehensive changes, institutions are encouraged to invite active translation professionals into the classroom to share their practical, up-to-date expertise.

Yes, technologies are essential to modern translation training, but they fundamentally change the translator’s cognitive processes and must be taught with a critical approach (Pym 2011b). Because tools like translation memories segment texts and distance translators from the overall communicative context, they risk reducing translation to a mechanical phrase-replacement process rather than an act of active communication. To counter this, Pym argues that educators should train students to work both with and against technology. Adaptation again comes to the fore here. Instead of dedicating extensive classroom hours to specific, rapidly outdated software, institutions should teach students how to learn independently and adapt to new tools. Furthermore, he emphasizes that although technology gets better at generating alternative translations, the core competence of selecting the single best solution still depends on human judgment and risk assessment.

The second cluster of works that bring technologies together with TT appeared in the second decade of the 2000s, starting with the article on changing translation skill sets (Pym 2013) discussed in the previous section. Addressing the localization and the training of localization professionals, also in line with his previous approaches to competence and technology, Pym (2014: 49) suggests that rather than merely supplying cheap labor to fulfill immediate market demands, academic training should intervene in the future of localization by equipping students with foundational technological skills together with a critical, overarching perspective on cross-cultural

communication. A later study by Pym and Torres-Simón (2016) has implications for various aspects of TT, including technologies and the profession. What particularly drew my attention in this study was the data collection method: Pym asked students in four different courses between 2012 and 2015 to formulate three questions they wanted to be answered about translation. One of the suggestions in this study, for instance, is that, instead of allowing students to perceive new technologies strictly as a threat to their future careers, educators can help students see these tools from a different perspective by integrating them into classroom teaching through hands-on experimentation, such as comparing the efficiency of translating from scratch versus post-editing machine translation (Pym and Torres-Simón 2016: 198).

In later years, Pym co-authored papers and monographs focusing on the teaching of translation technologies (Hao and Pym 2023a), trust in MT among translation students (Ayvazyan, Torres-Simón and Pym 2024) and the management and pedagogical use of GenAI in the translation classroom (Ayvazyan, Hao and Pym 2024; Pym and Hao 2025; Raigal-Aran et al. 2026).

Ayvazyan, Torres-Simón and Pym (2024) conducted a study with undergraduate translation students who completed two online tasks designed to assess their attitudes toward and performance with MT. Their aim is to open up “space where students can explore the technologies for themselves, reaching their own conclusions” (Ayvazyan, Torres-Simón and Pym 2024: 11). The findings indicate that while there is no unanimous stance on MT, students generally show strong resistance to it. Even though the participants clearly benefited from the technology, particularly in terms of translation speed, they maintained an unfavorable view that paradoxically led them to spend excessive time post-editing the machine output (Ayvazyan, Torres-Simón and Pym 2024: 25).

Raigal-Aran et al. (2026) further reported on a classroom experiment in which Spanish students evaluated and compared their own revisions of an AI-generated earthquake advisory text against a version refined by ChatGPT-4o. Although many students criticized the AI for taking “unjustifiable liberties,” roughly a third of the participants recognized that the generative AI provided valuable alternatives they had not originally considered (Raigan-Aral et al. 2026: 4). The activity demonstrated that GenAI can serve as a beneficial learning and supplementary tool, fostering a generally favorable attitude among students toward safely integrating these technologies to augment the human translation and revision process. The latter two studies are also significant in terms of methodology, as they allow students to experiment with MT to reach their own conclusions through hands-on engagement and reflection on their experiences.

Always interested in translation technologies, starting with the first discussions on CAT tools, Pym got interested in the main discussions of Translation Studies and TT in the last decade: MT and more generally GenAI in the translation classroom. In this group of works, a monograph stands out,

providing hands-on guidelines for teaching languages and translation in the age of GenAI and seeking an answer to the question: “How can we best engage with new learning technologies?” (Pym and Hao 2025: xviii). Rather than offering abstract warnings about technological disruption, the book moves directly into the classroom, presenting a wide range of activities through which students can find out what technologies can and cannot do. The reassuring message for trainers is that translators are not obsolete. Yet...

Market, employability and the sociology of the profession

Another thematic cluster places translation and TT within broader socio-economic contexts. With his background in sociology, Pym addresses the relationship between TT and the market, including the issues of disorder, employability, precarity and the inevitability of market change. His explicit sociological approach to training, particularly with reference to a sociology of TT, supports the view that educational models both reflect and affect professional realities.

Although many of his works address the evolution of the translation profession and education, particularly in relation to technology, we can find works specifically targeted at employability and the translation market, with references to training, starting from 1993. Over three decades ago, before MT became a central concern and before discussions of AI replacing the profession emerged, Pym (1993b, 1993c) argued against strictly tailoring TT to immediate, fluctuating market demands, advocating instead teaching translation as a broad, adaptable set of communication skills. Rather than producing mere “technicians” doing translation, Pym believes programs should use specialized fields as case studies to train critical “thinkers” or “cultural intermediaries” (*Sprachmittler*). These professionals are equipped to offer services that go beyond translation, such as editing, terminology work and company representation, enabling them to move beyond strict market logic and improve intercultural relations. His arguments remain relevant to training practices today, as the works discussed so far clearly anticipate today’s MT and GenAI discussions.

This perspective is taken up in Anthony Pym’s more recent work on the impact of NMT on the profession (Pym 2019b; Pym and Torres-Simón 2021). The year 2026 marks the tenth anniversary of Google’s announcement of NMT. This milestone calls for reflection on how the role of translators has changed, which Pym (2019b) summarizes through three main aspects: establishing their trustworthiness as authorities who post-edit MT output; highlighting the multiplicity of possible translation solutions and therefore defining translation as a problem-solving activity; and engaging in new professional domains, as “there is no real shortage of work to be done.”

In the same vein, Pym and Torres-Simón (2021: 53) emphasize that automation is reshaping the profession, arguing that “the more translators are

aware of automation and are prepared to work with it, the more they seek to have their multilingual interactive skills valued,” particularly in ensuring the trustworthiness of MT outputs. These works reach the same conclusion: TT should equip future professionals to move beyond mere translation by offering extended services, positioning themselves as trustworthy experts and acting as cultural mediators across sectors, especially since not all graduates are necessarily employed as translators.

That being said, Pym has co-authored several studies specifically addressing employability (Hao and Pym 2023b; Macreadie and Pym 2025; Pym and Macreadie 2026), whose findings align with his consistent arguments regarding TT and institutions. After a review of survey studies on the employability of translation graduates and their survey of graduates from the Chinese-English Master of Translation at the University of Melbourne, Hao and Pym (2023b) conclude that training institutions should acknowledge the fact that only about one third of graduates are employed in full-time jobs as translators or interpreters, and should therefore tailor their curricula to support a wider range of language-based occupations. Furthermore, programs should place greater emphasis on developing transferable skills, such as software evaluation, project management and interpersonal communication, rather than rigidly focusing on text-bound translation or specific software applications. Reflecting Pym’s long-standing argument (as discussed earlier) for teaching flexibility over narrow specialization, the authors argue that translation programs must abandon “disciplinary isolationism,” meaning they should stop distancing themselves from general language education, and instead train students to act as comprehensive language consultants who can offer a broad range of services well beyond translation.

Another employability study focuses on adverse selection, defined as an economic scenario in which consumers’ inability to independently judge product quality leads to lower prices, driving good professionals out of the market, risking a total market collapse (Pym and Macreadie 2026). Drawing on quantitative data from previous large-scale surveys of the Australian market and interviews with 35 stakeholders, Pym and Macreadie (2026) identify key signs of adverse selection in the Australian market. Although NAATI’s recertification system limits oversupply, certification alone cannot ensure stability. Low pay continues to drive quality down, which requires stronger labor protections. The authors also note that growing skepticism about quality demands more varied and visible quality signals. Rebuilding trust, Pym and Macreadie (2026) argue, means valuing both community-based “thick trust” and institutional “thin trust.” In the end, preventing market decline requires combined efforts at the policy, industry and practitioner levels.

Among the works grouped under this heading, one is particularly noteworthy for its attempt to define the sociology of translator training. (Pym 2020). In this paper, the argument is that how translators should be trained

depends on social factors, which Pym defines as the distribution of languages and the intensity of information flows. Understanding where a program is situated, who its students are and the types of professional activities it prepares them for enables trainers to draw selectively on diverse models and pedagogical approaches from different contexts. At the same time, it is important not to focus exclusively on the upper segments of the market, as if all graduates were destined for stable careers there. Such a view overlooks the broader social functions of translation and the range of contexts in which it functions. From a wider sense of social responsibility, TT should also engage more actively with additional-language education, where translation is often first encountered in simplified or reductive ways. Thus, TT should be understood as a context-dependent and evolving practice shaped by specific social needs, aspirations and technological developments, rather than as a fixed set of universal principles: “In each particular training situation, for each particular language pair, for each technological generation, we need to think anew” (Pym 2020: 15).

From a broader perspective, these issues are closely interconnected: the impact of changing technologies on the profession and, consequently, on TT; the need for flexibility in training in response to evolving technologies and market demands; and the growing importance of flexible employability.

And more...

The typology of translation solutions (Pym and Torres-Simón 2015; Pym 2016; Pym 2018b) differs from other typologies commonly found in Translation Studies in its explicit pedagogical orientation. Offering this typology as a “teaching aid” for solving translation problems in the classroom, Pym (2016) mentions that it is not a definitive list, but rather one that remains open to adaptation depending on students’ levels, language pairs and the specific focus of each lesson, i.e., students’ needs and expectations. This flexibility reflects the arguments above that TT should be responsive to context, learner needs and the changing nature of translation practice.

My own experience as a trainer has shown that knowing how to translate does not necessarily mean knowing how to teach translation. This is why Pym’s interest in translator trainers and especially their training (Pym 2001; Pym 2004b; Torres-Simón and Pym 2016; Hao and Pym 2022) is so important, as this issue is still underdiscussed in TT research. Pym (2001) approached the training of translation trainers as an experimental and practical necessity driven by rapid technological and economic changes in the professional translation market. Extending his arguments on student diversity, Pym identifies three groups requiring trainer training: language teachers entering translation, experienced teachers needing retraining on technology and professional translators needing pedagogical skills. Given their complementary strengths, he suggests that these groups can learn from each

other through active exchange, an approach reflected in the training programs he organized.

In the same vein, Hao and Pym (2022) conducted a study with two groups of learners, i.e., students trained in applied linguistics and those trained as translators. This study also highlights the merits of bringing together learners from different disciplinary backgrounds. While the initial design envisioned a reciprocal learning dynamic, where applied linguistics students would contribute expertise in language teaching and translation students would offer insights into translation practice, the actual composition and goals of the groups proved to be more heterogeneous than expected. Despite this mismatch, the findings suggest that well-articulated learning outcomes can effectively bridge the gap between using translation as a pedagogical tool in additional-language education and training translators and interpreters as professionals. More importantly, the study shows that such interdisciplinary settings can be justified through the complementary needs of the learner groups.

As reflected in some of his above works and specifically dedicated publications since the early 1990s, Pym repeatedly returns to the relationship between translation teaching and foreign language education (see, for example, Pym 1992; Pym 2005; Pym, Malkmjaer and Gutiérrez-Colón Plana 2013; Pym 2018a; Pym and Aylvazyan 2018). In these papers, Pym criticizes the academic and institutional divide between TT and foreign language teaching, noting that the translation field has historically sought to claim its independence and “essential superiority” (Pym 2018a) by establishing dedicated faculties and deliberately distancing itself from language education. This approach is clearly linked to his criticism of the implicit alignment between TT programs and the translation profession. He questions the assumption that translation programs should primarily produce professional translators, arguing that not all graduates enter the translation market and that translation itself is a communicative practice.

From this perspective, Pym calls for a more active dialogue between Translation Studies and language education. Emphasizing its communicative function enables translation to be integrated more meaningfully into language-learning contexts, moving beyond reductive or mechanical uses. At the same time, this perspective draws attention to the human, interpretive and decision-making aspects of translation, namely the aspects that have become more significant in an age dominated by AI and automation.

It is not possible to provide an overview of all his work here. What I find particularly noteworthy is Pym’s interest in relatively less discussed areas, such as work placement (Pym et al. 2014) and student exchange programs (Pym 1994), together with different aspects of classroom practices in translation teaching (see, for example, Pym 2009; Pym and Torres-Simón 2016; Pym 2023; Hao and Pym 2023a; Aylvazyan, Hao and Pym 2024), beyond the issues already discussed here.

Conclusion

This study outlines Anthony Pym's TT research over the past thirty-five years, offering a perspective on how TT has evolved. In his work, we observe a consistent effort to respond to technological and socio-economic changes not with rigid, "one-size-fits-all" curriculum structures, but with adaptable, context-dependent pedagogical strategies. At the core of Pym's philosophy is a redefined, minimalist model of translation competence: a two-fold process of generating viable target texts and selecting the most appropriate one. This specific problem-solving is particularly important today as we are discussing the inevitable effects of GenAI and NMT on the profession and translators are increasingly expected to select and adapt AI-generated outputs rather than produce texts from scratch.

Pym's pragmatic approach to technology, which argues that educators train students to work both "with" and "against" it, offers guidance for sustaining critical human judgment, risk assessment and the wider communicative context in which translation functions. Furthermore, Pym's sociological approach to TT challenges how educational institutions approach employability. Arguing against producing mere translators for a constantly changing market, he advocates for training "cultural intermediaries" equipped with transferable skills. This flexibility is vital today, as graduates face changing market conditions and assume various roles beyond pure translation.

Declaration of AI use

In preparing this manuscript, the author used ChatGPT (GPT-5.5) and Grammarly for drafting and language editing. All AI-assisted content was reviewed, edited and verified by the author, who takes full responsibility for the accuracy and integrity of the final text.

References and a(n) (almost) complete list of Anthony Pym's works on translator training

- Ayvazyan, Nune, Ester Torres-Simón, and Anthony Pym. 2024. "Translation students' trust in machine translation: Too good to be true?" In M. C. Odacıoğlu (ed) *Revolutionizing Translation Studies. Synthesizing Translation with AI and IT Innovation*. Istanbul: Aktif. 9-30. <https://doi.org/10.17613/ss86-h148>
- Ayvazyan, Nune, Yu Hao, and Anthony Pym. 2024. "Things to do in the translation class when technologies change. The case of generative AI". In Y. Peng, H. Huang, and D. Li (eds.) *New Advances in Translation and*

- Interpreting Technology: Theories, Applications and Training*. Singapore: Springer. https://doi.org/10.1007/978-981-97-2958-6_11
- Caminade, M., and Anthony Pym. 1998. "Translator-training institutions". In M. Baker (ed) *Encyclopedia of Translation Studies*. London and New York: Routledge. 280-285.
- Fallada-Pouget, Carmina, and Anthony Pym. 2000. "Learner-centered distance education: A literature review of the 'no difference' debate". In F. Austerlühl and D. Apollon (eds) *Humanities Education and the Challenge of E-Learning*. Bergen: The HIT Centre Publications, University of Bergen. 41-47.
- Hao Yu, and Anthony Pym. (2022). "Teaching how to teach translation: Tribulations of a tandem-learning model". *Perspectives: Studies in Translatology* 30(2), 275-291. <https://doi.org/10.1080/0907676X.2021.1913197>
- Hao Yu, and Anthony Pym. 2023a. "Choosing effective teaching methods for translation technology classrooms: Teachers' perspectives". *Forum* 21(2): 190-202. <https://doi.org/10.1075/forum.22017.hao>
- Hao Yu, and Anthony Pym. 2023b. "Where do translation students go? A study of the employment and mobility of Master graduates". *The Interpreter and Translator Trainer* 17(2), 211-229. <https://doi.org/10.1080/1750399X.2022.2084595>
- Macreadie, Rachel, and Anthony Pym. 2025, April. "Why interpreters and translators say they are likely to leave the sector". *InTouch* 33(1), 15.
- Macreadie, Rachel, Kadija Bouyzourn, Anthony Pym, and Reine Meylaerts. 2025. "Thick and Thin Trust in Translated Cross-Cultural Vaccination Messaging". *Translation Studies* 18(2), 265-283. <https://doi.org/10.1080/14781700.2025.2516000>
- Pym, Anthony. 1991. "A definition of translational competence, applied to the teaching of translation". In M. Jovanovic (ed.) *Translation: A Creative Profession: 12th World Congress of FIT*. Proceedings. Belgrade: Prevodilac. 541-546.
- Pym, Anthony. 1992. "Translation error analysis and the interface with language teaching". In C. Dollerup and A. Loddegaard (eds) *The Teaching of Translation*. Amsterdam: John Benjamins. 279-288. <https://doi.org/10.1075/z.56.42pym>
- Pym, Anthony. 1993a. *Epistemological Problems in Translation and Its Teaching*. Calaceite: Caminade.
- Pym, Anthony. 1993b. "On the market as a factor in the training of translators". *Koinè* 3: 109-121.
- Pym, Anthony. 1993c. "Market-based arguments against the market as a direct factor in the training of translators". *Journal of the Association of Slovak Anglicists* 2(2): 7-11.

- Pym, Anthony. 1994. "Student exchange programmes and translator training: Three economic principles". *Perspectives: Studies in Translatology*, 2(1): 41-50. <https://doi.org/10.1080/0907676X.1994.9961221>
- Pym, Anthony. 1996. "Ideologies of the expert in discourses on translator training". In M. Snell-Hornby and Y. Gambier (eds) *Problems and Trends in the Teaching of Interpreting and Translation*. Misano: Istituto San Pellegrino. 139-149.
- Pym, Anthony. 1999. "Localizing localization in translator-training curricula". *Linguistica Antverpiensia* 33: 127-137.
- Pym, Anthony. 2000a. "Innovation in translator and interpreter training: Report on an on-line symposium". *Across Languages and Cultures* 1(2): 209-211. <https://doi.org/10.1556/Acr.1.2000.2.5>
- Pym, Anthony. 2000b. "On the distance in distance learning". In F. Austerlühl and D. Apollon (eds) *Humanities Education and the Challenge of E-Learning*. Bergen: The HIT Centre Publications, University of Bergen. 55-62.
- Pym, Anthony. 2001. "Trial, error and experimentation in the training of translation teachers". *Traducción & Comunicación* 2: 73-90.
- Pym, Anthony. 2002a. "E-learning and translator training". In *Current Status of Translation Education: International Conference on Translation Studies*. Seoul: Sookmyung Women's University. 3-36.
- Pym, Anthony. 2002b. "Training language service providers: Local knowledge in institutional contexts". In B. Maia, J. Haller and M. Ulrych (eds) *Training the Language Services Provider for the New Millennium*. Porto: Universidade do Porto. 21-30.
- Pym, Anthony. 2003. "Redefining translation competence in an electronic age". *Meta* 48(4): 481-497. <https://doi.org/10.7202/008533ar>.
- Pym, Anthony. 2004a. "E-learning in translator training". In L. Rega and M. Magris (eds) *Übersetzen in der Fachkommunikation – Comunicazione specialistica e traduzione*. Tübingen: Gunter Narr. 161-173.
- Pym, Anthony. 2004b. "On the training of translation teachers: Report on a consortium". In *A Profissionalização do Tradutor*. Lisbon: Fundação para a Ciência e a Tecnologia, União Latina. 53-58.
- Pym, Anthony. 2005. Training translators – ten recurrent naiveties. *Translating Today* 2: 3-6.
- Pym, Anthony. 2006. "Asymmetries in the teaching of translation technology". In A. Pym, A. Perekrestenko and B. Starink (eds) *Translation Technology and Its Teaching (with much mention of localization)*. Tarragona: Intercultural Studies Group. 113-124.
- Pym, Anthony. 2008a. "Translation technology and training for intercultural dialogue: What to do when your translation memory won't talk with you". In R. Dimitriu and K.-H. Freigang (eds) *Translation Technology in Translation Classes*. Iași: Institutul European. 12-27.

- Pym, Anthony. 2008b. "Professional corpora: teaching strategies for work with online documentation, translation memories and content management". *Chinese Translators' Journal* 29(2): 41-45.
- Pym, Anthony. 2009. "Using process studies in translator training. Self-discovery through lousy experiments". In S. Göpferich, F. Alves and I. M. Mees (eds) *Methodology, Technology and Innovation in Translation Process Research*. Copenhagen: Samfundslitteratur. 135-156.
- Pym, Anthony. 2011a. "Training translators". In K. Malmkjær and K. Windle (eds) *The Oxford Handbook of Translation Studies*. Oxford: Oxford University Press. 475-489.
<https://doi.org/10.1093/oxfordhb/9780199239306.013.0032>
- Pym, Anthony. 2011b. "What technology does to translating". *Translation & Interpreting* 3(1): 1-9.
- Pym, Anthony. 2013. "Translation skill-sets in a machine-translation age". *Meta* 58(3): 487-503. <https://doi.org/10.7202/1025047ar>
- Pym, Anthony. 2014. "Localization, training, and instrumentalization". In E. Torres-Simón and D. Orrego-Carmona (eds) *Translation Research Projects 5*. Tarragona: Intercultural Studies Group. 37-50.
- Pym, Anthony. 2015. "Translating as risk management". *Journal of Pragmatics* 85: 67-80. <https://doi.org/10.1016/j.pragma.2015.06.010>
- Pym, Anthony. 2016. *Translation Solutions for Many Languages – Histories of a Flawed Dream*. London: Bloomsbury.
- Pym, Anthony. 2018a. "Where Translation Studies lost the plot: Relations with language teaching". *Translation and Translanguaging in Multilingual Contexts* 4(2): 203-222.
<https://doi.org/10.1075/ttmc.00010.pym>
- Pym, Anthony. 2018b. "A typology of translation solutions". *Journal of Specialised Translation*, 30: 41-65.
- Pym, Anthony. 2019a. "Teaching translation in a multilingual practice class". In D. Sawyer, F. Austerlühl and V. Enríquez Raído (eds) *Translator Education: The Evolving Curriculum*. Amsterdam and Philadelphia: John Benjamins. 75-79.
- Pym, Anthony. 2019b. "How neural machine translation might change the work of professional translators". In H. Che Omar et al. (eds) *Translation in the Digital Era. Proceedings of the 17th International Conference on Translation*. Kuala Lumpur: Universiti Sains Islam Malaysia. 1-13.
- Pym, Anthony. 2019c. "Merkel Responds to Reem: What we should be training mediators for". In A. Sohár, I. Limpár and D. Galambos (eds) *Getting Translated*. Budapest: Tinta Könyvkiadó. 9-18.
- Pym, Anthony. 2020. "For a sociology of translator training". In J. Zhao, D. Li and L. Tian (eds) *Translation Education. A Tribute to the Establishment of the World Interpreter and Translator Training Association (WITTA)*. Singapore: Springer Nature. 9-19.

- Pym, Anthony. 2021. "Cooperation, risk, trust. A restatement of translator ethics". *Stridon. Journal of Studies in Translation & Interpreting* 2(1): 5-24. <https://doi.org/10.4312/stridon.1.2.5-24>
- Pym, Anthony. 2023. "Active translation literacy in the literature class". *PMLA* 138(3): 819-823. <https://doi.org/10.1632/S0030812923000718>
- Pym, Anthony, and Ayvazyan, N. 2018. "Linguistics, translation and interpreting in foreign-language teaching contexts". In K. Malmkjær (ed) *The Routledge Handbook of Translation and Linguistics*. London and New York: Routledge. 393-407.
- Pym, Anthony, Bei Hu, Maria Karidakis, John Hajek, Robyn Woodward Kron, and Riccardo Amorati. 2023. "Community trust in translations of official COVID-19 communications in Australia: An ethical dilemma between academics and news media". In P. Blumczynski and S. Wilson (eds) *The Languages of COVID-19. Translational and Multilingual Perspectives on Global Healthcare*. New York: Routledge. 110-127. <https://dx.doi.org/10.4324/9781003267843-10>
- Pym, Anthony, David Orrego-Carmona, and Ester Torres-Simón. 2016. "Status and technology in the professionalization of translators. Market disorder and the return of hierarchy". *The Journal of Specialised Translation* 25: 33-53.
- Pym, Anthony, and Ester Torres-Simón. 2015. "The pedagogical value of translation solution types". *Perspectives: Studies in Translatology* 23(1), 89-106. <https://doi.org/10.1080/0907676X.2014.928334>
- Pym, Anthony, and Ester Torres-Simón. 2016. "Designing a course in Translation Studies to respond to students' questions". *The Interpreter and Translator Trainer* 10(2), 183-203. <https://doi.org/10.1080/1750399X.2016.1198179>
- Pym, Anthony, and Ester Torres-Simón. 2021. "Is automation changing the translation profession?". *International Journal of the Sociology of Language* 270: 39-57. <https://doi.org/10.1515/ijsl-2020-0015>
- Pym, Anthony, and José Ramón Biau Gil. 2006. "Technology and translation: A pedagogical overview". In A. Pym, A. Perekrestenko and B. Starink (eds) *Translation Technology and Its Teaching (with much mention of localization)*. Tarragona: Intercultural Studies Group. 5-20.
- Pym, Anthony, and Yu Hao. 2025. *How to Augment Language Skills. Generative AI and Machine Translation in Language Learning and Translator Training*. London: Routledge.
- Pym, Anthony, and Rachel Macreadie. 2026. "Translator employment and the risk of market collapse. News from Australia". *Translation & Interpreting*, 18(1): 1-18. <https://doi.org/10.12807/ti.118201.2026.a01>
- Pym, Anthony, Gabriel González Núñez, Marta Miquel-Iriarte, Sara Ramos Pinto, Carlos S. C. Teixeira, and Wine Tesseur. 2014. "Work placements in doctoral research training in the humanities: Eight cases from

- Translation Studies". *Across Languages and Cultures*, 15(1), 1-23. <https://doi.org/10.1556/Acr.15.2014.1.1>
- Pym, Anthony, Kirsten Malmkjaer, and Maria del Mar Gutiérrez-Colón Plana. 2013. *Translation and Language Learning: The Role of Translation in the Teaching of Languages in the European Union. A Study*. Luxembourg: European Commission.
- Raigal-Aran, Judith, Nune Ayvazyan, Yu Hao, Y., and Anthony Pym. 2026. *Generative AI in the Translation Revision Class. Technical Report on the Spanish Activity*. <https://doi.org/10.13140/RG.2.2.33219.26406>
- Torres-Simón, Ester, and Anthony Pym. 2016. "The professional backgrounds of translation scholars: Report on a survey". *Target* 28(1): 110-131. <https://doi.org/10.1075/target.28.1.05tor>
- Torres-Simón, Ester, and Anthony Pym. 2019. "European Masters in translation: A comparative study". In D. Sawyer, F. Austerlühl and V. Enríquez Raído (eds) *Translator Education: The Evolving Curriculum*. Amsterdam and Philadelphia: John Benjamins. 319-340.

Watching the process unfold: using API-based interaction data to understand student use of AI in translation tasks

JUDITH RAIGAL-ARAN and NUNE AYVAZYAN

Universitat Rovira i Virgili, Spain

This article describes a methodology to automatically create translation memories for subtitling, using translated books adapted into films and recognizing extra-linguistic markers to differentiate character interventions from narration. This methodology includes the automatic identification, extraction and alignment of the dialogues. The aligned bi-texts served as translation memories in the subtitling of the adapted films. Results show an overall 95% extraction rate for English dialogues and 85% for Spanish dialogues. Alignment showed an accuracy of 90%. Results for the translation memory performance showed that hits between 70% and 100% matches accounted for 15% of the corpus. The results reinforce the claim that dialogues in books can be used as reference material for the translation of subtitles.

Keywords: audiovisual translation, translation memories, subtitling, natural language processing

Introduction

Anthony Pym has made a sustained and influential contribution to both translator education and the pedagogical use of translation as a means of learning languages. Over the years, his work has helped shape how translation is taught, not only as a professional skill but also as a valuable communicative practice within language learning (Pym et al. 2013; Pym and Ayvazyan 2017). Across his work on translator training, Pym has consistently argued that translation skills should be complemented with wider translator competence, including risk management, critical reflection and trust (Pym 2015; 2025). His research on translation and language learning has further highlighted how translation tasks can be used to develop language awareness and intercultural competence, challenging rigid binary separations between “language

teaching” and “translation teaching” (Pym et al. 2013) and calling for empirically grounded accounts of classroom practice (Pym 2011). Pym has also underscored the need to integrate emerging and evolving technologies into pedagogy in ways that are both exploratory and critically informed (Ayvazyan et al. 2024). The rapid diffusion of large language models (LLMs) in translation studies extends these concerns into a new technological moment. Pym’s recent work with Yu Hao on generative AI (2025) explicitly frames such tools as potential means of augmenting language skills while practicing translation.

The integration of machine translation and, more recently, LLMs into both translator training and language learning has renewed longstanding interrogations about questions such as what competences should be cultivated, how technology mediates learning and what role risk, trust and critical reflection play in the training of translators, questions that were already central in the EMT 2022 competence framework.

These questions have become increasingly urgent, yet empirical research into how students actually use AI during translation tasks remains limited. Much of the existing evidence on generative AI use in translator training is based on self-reported data such as surveys, interviews or post-hoc reflections (*cf.* Almahasses et al. 2024; Belhassen and Hamda 2025; Bouyzourn and Birch 2025) which, while valuable, cannot capture the granular, process-oriented dimension of AI-mediated practice: how students formulate queries, revise prompts, selectively accept or reject suggestions and negotiate with the system across iterative exchanges. This gap between students’ accounts of their AI use and the actual texture of that use constitutes a significant methodological challenge for the field.

This study responds to this challenge by reporting on an experiment designed to move beyond self-report studies towards the systematic analysis of interaction logs. By situating our study within the concerns that have shaped Pym’s research trajectory—technology, risk and trust applied to translator training and language learning—we seek to contribute to an empirically grounded understanding of how generative AI is being incorporated, resisted and renegotiated in translator and language training classrooms.

Literature review

Process-oriented research on AI-assisted translation tools such as Google Translate and DeepL has historically employed methodologies such as eye tracking, keylogging to capture fine-grained data on user behavior, as well as think-aloud protocols (Muñoz and Rojo 2025). These approaches make it possible to examine attention patterns, revision sequences and decision-making processes as they unfold in real time. Eye-tracking studies reveal how

users allocate visual attention across source texts, machine output and external resources, keylogging provides detailed records of text production and editing, while think-aloud protocols offer access to participants' verbalized reasoning. Together, these methodologies try to shed light on the cognitive processes underlying translation decisions, while these remain largely inaccessible and can only be inferred indirectly through such observable data.

The analysis of student prompting strategies represents a new avenue for process-oriented research; as such, relatively few studies have examined actual interactions through prompts. Zhang et al. (2025) studied student prompting strategies for five translation-related tasks: understanding, transfer, documentation, revision and analysis. The data revealed that the students had used generative AI mostly for transfer and revision. The students also used the tool to obtain explanations of domain-specific terms, proper names and background information. With regard to prompting behavior, approximately half of the students produced a single prompt, whereas the remaining participants engaged in multi-turn interactions with the system. Further, most prompts were expressed in the imperative but also included requests and suggestions. Similarly, Su et al. (2026) found that students tend to adopt direct prompting strategies when interacting with general-purpose generative AI tools, often requesting complete answers or executing straightforward commands rather than engaging in more exploratory or reflective forms of interaction.

Some studies have tried to examine trust in generative AI through surveys. Bouyzourn and Birch (2025) surveyed 115 university students and found generally high levels of trust in ChatGPT. Frequent use increased trust, while greater technical understanding reduced it. Translation was considered moderately reliable and only 8.7% of participants expressed negative views of the tool. Almahasees et al. (2024) surveyed 102 English students and found generally positive perceptions of ChatGPT across several dimensions. Students reported high trust in its translations and believed they preserved source meaning. Many participants also expressed trust in the system's ability to maintain the confidentiality of their translated work. Belhassen and Hamda (2025) surveyed 150 students and found high but conditional trust in AI-generated translations. Most participants considered outputs accurate (83.3%) and expressed trust in AI tools (66.6%). However, a third noted limitations in handling complex tasks requiring creativity, cultural awareness, or deeper understanding.

To our knowledge, no study has yet examined prompt formulation behavior among translation students working specifically with legal texts. This chapter seeks to address this gap.

Methodology

The data reported in this article were collected during a single in-class session in November 2025 as part of *Direct Translation I*, an elective course taken by third- and fourth-year students enrolled in a bachelor's degree in English. Although 19 students were present on the day of the experiment, the analysis includes only the 16 participants who carried out the task successfully (completed both pre and post questionnaires). The study received clearance from the Ethics Committee for Research on People, Society and the Environment at Universitat Rovira i Virgili, with the approval code CEIPSA-2024-PR-0012.

The course is one of two translation courses in the degree, alongside the mandatory third-year course *Introduction to English Translation*. Students are generally expected to complete the mandatory course where foundational translation skills are taught *before* enrolling in the elective. So, the students had some prior training in translation. The elective course *Direct Translation I* incorporates components of specialized translation, including technical, academic, economic and legal domains. In the course, students are required to translate from English (their L2/L3) into their preferred L1, Spanish or Catalan. Legal translation is generally regarded by students as particularly demanding due to its reliance on system-specific, dense and often non-transparent terminology and the potential consequences of interpretative errors. For this reason, a legal text was selected for the experiment, as it was expected to elicit a greater number and variety of student queries related to meaning negotiation, term selection and reformulation strategies when interacting with the GPT API.

The legal text provided was a standard force majeure clause in English (189 words), part of an agreement. It contained several translation problems primarily related to terminology (e.g. “act of God”), expressions that are inconsistently interpreted across legal systems (e.g. “best efforts”) as well as references to applicable law (e.g. “Cal. Com. Code § 2-615”). The full text can be consulted in Section 4, Results.

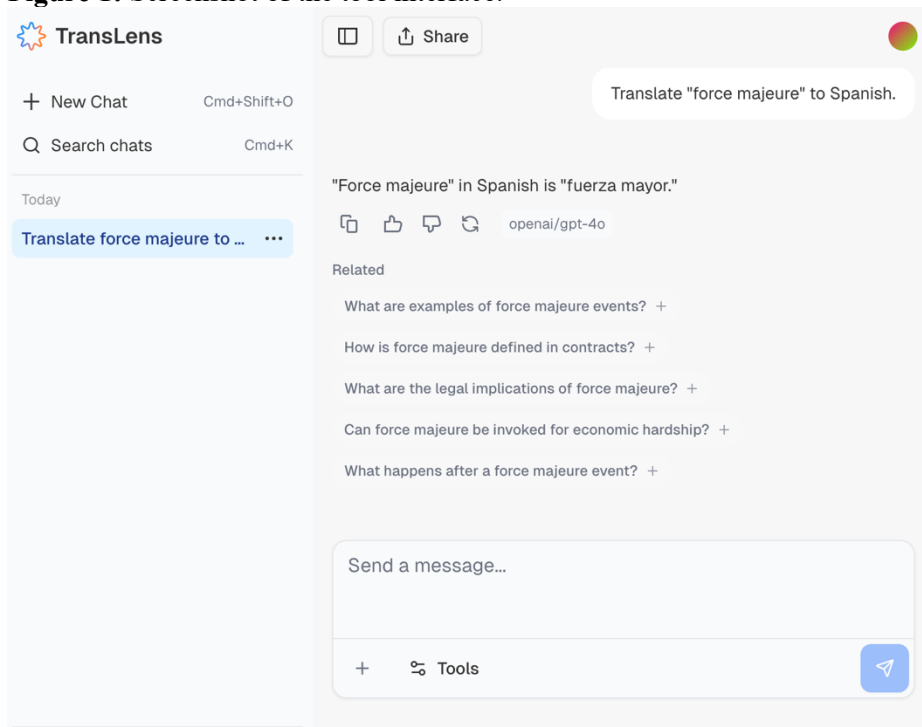
The session was designed as a timed individual translation activity which lasted 70 minutes, structured around four sequential steps that were explained to the students prior to the experiment. The four steps consisted of a pre-task questionnaire, translation, a post-task questionnaire and a final class discussion of the experiment. The study aimed to answer the following research questions:

- Q1. What interaction and revision patterns can be observed in students' use of generative AI?
- Q2. How do students formulate prompts when using generative AI for legal translation? and
- Q3. Does students' trust in AI-generated translations increase after the task?

The students began by completing a pre-task questionnaire designed to capture their prior experience with AI translation tools and their expectations regarding the task ahead. They were then asked to translate a 189-word legal text from English into either Spanish or Catalan, both working languages of our students. To contextualize the translation task, students were provided with a translation brief: a company based in Tarragona was preparing to sign a service agreement with a Californian company and required a translation of the document in order to assess several key clauses before signing. This framing was intended to situate the task within a recognizable professional scenario and to provide a communicative purpose for the translation.

For the translation step, the students were encouraged, though not required, to use GPT-4o as part of their workflow. Access to the tool was provided through a dedicated account linked to a paid API subscription set up specifically for this activity. Each student was assigned an individual anonymized username and password, a configuration that served a dual purpose: it allowed students to access a version of the tool with greater capacity than the free tier and it enabled the researchers to retrieve the complete interaction logs associated with each account for subsequent analysis. A screenshot of the tool interface can be seen in Figure 1.

Figure 1: Screenshot of the tool interface.



The tool, designed to resemble the interface of ChatGPT (or any chat-based assistant), was specifically developed for this activity and enabled the systematic recording of all interactions between students and GPT-4o. These interactions were subsequently made available for retrieval in the form of an Excel file. The dataset includes columns for anonymized user identifiers (created expressly for the activity), conversation id, message id, timestamps corresponding to each message (day and hour), message authorship (user or assistant) and the message content itself.

The students were informed that the session data would be recorded and used for research purposes and were instructed not to include any personal information in their queries. They were also free to consult any other resources they considered relevant, meaning that AI use was positioned as one option among several rather than as a mandatory tool. Screen activity was not recorded; instead, students self-reported their use of additional resources in the post-task questionnaire.

Upon completing their translation, the students responded to a post-task questionnaire. This second questionnaire aimed at gathering post-task reflections on their use of the tool, the difficulties encountered and their assessment of the translation they had produced. The class then finished with a 20-minute discussion where the students discussed the activity.

This design allowed to triangulate four different data sources: the pre- and post-task questionnaires, the final translations submitted by students, the full logs of student-AI interactions captured through the API and the class discussion. In this article, we report on the data we obtained from the student-AI interactions through the API but also address key questions on trust in the pre- and post-task questionnaires. Due to the space constraints of the present article, the remaining data will be presented in a separate, related article.

Results

This section presents the results obtained from 16 students. Within the dataset generated by the GPT API interface, the participants were labeled as “users.” For consistency reasons, the participants are therefore identified as “users” followed by their numerical identifier as assigned by the GPT API, while the term “students” is employed in the narrative when referring to the cohort more generally.

Interaction counts, time-on-tool and time-on-task

Interaction data was extracted from the conversation logs. We computed each student’s number of user-initiated messages (i.e., prompts sent to GPT), total messages exchanged, session start time (first prompt) and end time (last GPT response).

Combining student-initiated prompts and GPT responses, the total number of messages exchanged per student ranged from 2 to 24 ($M = 8.12$ exchanges), with a grand total of 132 messages across the full sample (see Table 1). The distribution was markedly asymmetric and polarized into three clusters. Eight students (50%) exchanged four messages or fewer, equivalent to one or two conversational turns. This suggests a largely transactional use of the tool. A middle cluster of four students (user12, user13, user14, user15) exchanged six messages each, consistent with a brief but slightly more iterative engagement. At the upper extreme, two students engaged in 16 interaction turns (user11 and user16) and two students (user2 and user4) accounted for 48 of the 132 total messages (36.36% of the entire sample's conversational output), with 24 messages each.

Table 1: Interaction counts, time-on-tool and time-on-task

Student	Interactions (n)	GPT session duration	Total task time
user1	6	11m 25s	43m 41s
user2	24	21 m 41s	51m 12s
user3	6	11m 10s	36m 6s
user4	24	10m 53s	40m 16s
user5	6	2m 7s	32m 53s
user6	2	0m 7s	36m 29s
user7	6	4m 40s	35m 3s
user8	2	0m 10s	39m 13s
user9	2	0m 8s	35m 38s
user10	4	23m 13s	(no pre-task)
user11	16	7m 20s	38m 56s
user12	4	22m 44s	33m 47s
user13	4	0m 11s	32m 30s
user14	4	2m 42s	49m 2s
user15	4	2m 58s	33m 52s
user16	16	8m 53s	21m 9s
Total	132	-	-
Mean	8.12	8m 9s	37m 34s

The total interaction time of each participant with the tool was calculated using the timestamp of the participant's first input and the timestamp of the

assistant's last response. This ranges from as little as seven seconds (user6) to over twenty-three minutes (user10), with a mean of 8 minutes and 9 seconds.

Total task duration, which is the time spent on the pre-task questionnaire, the translation task itself (including GPT interaction) and the post-task questionnaire, measured from the start of the pre-task questionnaire to the completion of the post-task questionnaire, ranged from 21 to 51 minutes (M = 37 minutes and 34 seconds). All the participants completed all the stages of the experiment, except for user10, who did not complete the pre-task questionnaire.

What is interesting here is that the time spent on the GPT API does not seem to reflect the number of messages: user2 spends 21m 41s on the tool to interchange 24 messages, while user4 only needs 10m 53s for the same number of interactions. On the other hand, user10 and user12 use roughly the same amount of time on the tool as user2, but with only four interactions each. Thus, time spent on the tool does not necessarily indicate the amount of interaction. Further, the quality of the final text cannot be inferred from interaction volume alone. Interaction metrics must be interpreted in relation to qualitative aspects of use, such as the complexity of prompts and the extent of revision, together with the final text quality.

The time logs also show that several participants (e.g., user6, user8, user9) interact with the system for only a few seconds, yet their total task time remains comparable to that of other students. This suggests that a substantial portion of the translation process takes place outside the GPT interface, which may include reading, revising, or evaluating the generated output.

Prompt patterns and query strategies

Half of the students (n = 8) wrote their prompts in Spanish, their L1 and the target language of the translation, while the other half chose English, the source language of the text (see Table 2). This split suggests that students did not treat prompt language as a principled choice aligned with the translational direction of the task.

In terms of syntactic form, imperative (e.g., "Translate this text into spanish" [sic], as expressed by user7) constructions dominated (n = 9, 56.3%), while only two students framed their initial query as a question (e.g., "Can you translate [...]"), as expressed by user11). The very limited use of interrogative forms further reinforces this tendency, as few students framed the interaction as a request for assistance or dialogue. Instead, prompting was largely reduced to issuing instructions, which may constrain opportunities for iterative refinement or exploratory engagement with the system. The register was markedly direct: only one student greeted the system ("hey chat", user1) and only one used "please" (user3), which indicates that most participants approached the tool as a utility rather than a conversational agent.

All 16 students who used GPT requested a translation of the full text at some point during the session. Most included the source text in their initial prompt ($n = 15$, 93.75%) and approximately two thirds specified the target language ($n = 10$, 62.5%) or provided a translation brief ($n = 9$, 56.25%), which was actually the commission given to them in the activity instructions. Very few mentioned the source language explicitly ($n = 2$, 12.5%) and only five students added instructions beyond a simple directive to translate. For example: “Traduce el siguiente texto jurídico” [translate this legal text] (user11), “Traduce este fragmento de un contrato, en el que se especifican las causas de fuerza mayor y como [sic] esto afecta a dicho contrato, del inglés al español” [Translate this excerpt from a contract, which specifies force majeure circumstances and how they affect the contract, from English into Spanish] (user14).

Table 2: Initial prompt characteristics

Variable	Category	n	%
Language of prompts			
	Spanish (L1)	8	50%
	English (L2)	8	50%
Prompt form			
	Imperative	9	56.3%
	Question	2	12.5%
	Teacher-provided default instructions (unmodified)	5	31.3%
Initial prompt content			
	Requested translation of the full text	15	93.75%
	Included the source text	15	93.75%
	Included the translation brief	9	56.25%
	Specified target language	10	62.5%
	Specified source language	2	12.5%
	Added other instructions	5	31.25%

To provide a clearer illustration of a fully developed prompt, the following example shows how one student formulated their request (in italics) by asking the tool to translate the text and by including both the source text and the brief outlined in the instructions:

Please help me with the Spanish translation of the following text. Take into account possible specific terms and keep the same tone and language: "6.1 In case of earthquake, typhoon, flood, fire, war or unavoidable force majeure events (including but not limited to act of God, strike, riot, act of war or outbreaks of infectious diseases), and

thereby causing direct impact on the performance of this Agreement or this Agreement cannot be performed according to the terms agreed, the party who encounters the aforesaid force majeure event shall immediately inform the other party, provide written report on details of force majeure within 5 business days after the occurrence of force majeure event, and submit valid supporting documents. Based on the event's degree of impact on agreement performance, all parties may decide whether or not to exempt from performing the obligations of this Agreement, or delay the performance of this Agreement. The performance affected by such force majeure event may be delayed or excused under Cal. Com. Code § 2-615. Neither party may propose a claim for compensation for the loss caused by force majeure. In addition, the parties shall use best efforts to mitigate the impact of the force majeure event and to resume performance of this Agreement as soon as reasonably possible. Pay special attention to these guidelines: Descripción del encargo. Una empresa con sede en Tarragona tiene previsto firmar un contrato de intercambio de datos y servicios con una empresa californiana. Antes de la firma, la empresa de Tarragona ha identificado dudas significativas sobre varias cláusulas clave, como la responsabilidad, las condiciones de resolución de disputas, la protección de datos personales, la fuerza mayor y la ley aplicable. Para tomar una decisión informada y minimizar riesgos, la empresa te solicita una traducción del documento.

The student's prompt demonstrates a structured strategy that combines task definition, contextualization and guidance. The request includes both the full source text and the translation brief, which situates the task within a professional context and may support more appropriate output. The reference to tone and terminology indicates awareness of domain-specific challenges. However, the instructions remain somewhat general, as no concrete constraints or priorities are specified other than the reference to the brief.

Reformulation behavior (see Table 3) was less common but qualitatively significant. Three students asked GPT to revise the full translation (user2, user7 and user11), though only user11 did so specifying that it was a legal text. Three others requested reformulation at the level of individual terms or fragments (user4, user7 and user11) and four students introduced targeted changes to punctuation or word choice. Perhaps most noteworthy is the behavior of three students who explicitly challenged GPT's translation decisions. They questioned, for instance, the capitalization of "Agreement" or the choice of a specific equivalent (translation of "acts of God" as "actos de Dios"). One of the students (user4) asked GPT to implement each change directly in the chat, effectively using it as a text editor and delegating the writing process itself, thereby turning GPT into an active co-editor.

Table 3: Reformulation behavior

Types of revision requests	n	%
----------------------------	---	---

Requested reformulation of full translation	3	18.8%
Reformulation included specific instructions	1	6.3%
Requested reformulation of a term or fragment	3	18.8%
Requested specific changes (punctuation, terminology, etc.)	4	25.0%
Challenged GPT's translation decisions	3	18.8%
Introduced changes from an external source	2	12.5%

When it comes to terminology-specific queries (see Table 4), seven students did at least one terminology-related query (user12 and user14 did one, user10 and user 13 did two, user4 did three, user16 did 4 and user11 did six). Across these seven prompts, the users engaged GPT as a combined translation, revision and editing tool, asking for legal terminology equivalents (e.g., “performance”, “party” and references like “Cal. Com. Code § 2-615”), questioning translation choices and consistency (especially capitalization and term selection) and issuing direct instructions to modify the text (such as removing commas, replacing words, or inserting translated segments).

Table 4: Terminology queries

Presence of query	n	%
At least one terminology query	7	43.8%
No terminology queries	9	56.3%

A comparison with Table 1 shows that, unsurprisingly, most students with more interactions also produced a greater number of queries: user4 with a total of 24 interactions and user11 and user16 with 16 interactions each. Interestingly, user2, who also engaged in 24 interactions and spent roughly double the time spent by user4 on the GPT API, had no terminology queries. These findings suggest that while a higher number of interactions tend to correlate with a greater number of queries, this relationship is not straightforward. The case of user2 demonstrates that extensive interaction and longer time spent on the tool do not necessarily translate into deeper or more focused engagement with specific translation challenges such as terminology.

Acceptance and modification of AI-generated output

Table 5 presents the number of modifications (if any) per student. Of the 16 students, nine (56.25%) made no changes to the text provided by GPT and submitted the AI-generated text without any modifications (100% similarity). The seven remaining students did edit the text (between 7 and 46 changes). It is noteworthy that during the post-task discussion, one student reported consulting GPT but ultimately relying on a machine translation tool to complete the translation.

The students who revised the text made such changes as lexical substitutions (selecting different Spanish equivalents for English terms), structural reformulations and, in some cases, additions of words not present in the GPT output. The student with the highest modification rate (user13, 29.1%) only had a single query but consistently reformulated terminology and restructured several clauses, suggesting an active and critical post-editing process that unfolded outside the GPT interface. At the other extreme, five students submitted only a single query to the tool and made no changes.

Table 5: Modification of AI-generated output

User	Words in GPT output	Words in student text	Changed words	% changes	Interactions (n)	GPT session duration
User1	235	238	45	19.1	6	11m 25s
User2	231	231	17	7.4	24	21 m 41s
User3	234	234	0	0	6	11m 10s
User4	242	242	0	0	24	10m 53s
User5	193	239	46	23.8	6	2m 7s
User6	241	241	0	0	2	0m 7s
User7	219	219	0	0	6	4m 40s
User8	229	229	0	0	2	0m 10s
User9	227	227	0	0	2	0m 8s
User10	240	239	32	13.3	4	23m 13s
User11	238	238	0	0	16	7m 20s
User12	238	237	7	2.9	4	22m 44s
User13	220	236	64	29.1	4	0m 11s
User14	241	247	26	10.8	4	2m 42s
User15	241	241	0	0	4	2m 58s
User16	240	240	0	0	16	8m 53s

As can be seen in Table 5, there is no consistent relationship between interaction metrics and revision behavior. For example, user2 engaged in 24 interactions and spent over 21 minutes on the tool but introduced only limited changes (7.4%), while user13 made the highest number of changes (29.1%) with only four interactions and almost no recorded time on the GPT interface. Similarly, several users with multiple interactions (e.g., users 4, 11 and 16) made no changes at all. In the case of these users, their interactions dealt with local adjustments to punctuation, capitalization (“Acuerdo”), specific terms (e.g. “días laborables” vs. “días hábiles”) and the explicit reference to the “*Cal. Com. Code § 2-615.*” User14 similarly concentrated on micro-level choices such as the translation of “party” as “partes” and the phrasing of “as

soon as reasonably possible”, without prompting the system for substantially different versions of the clause and not changing the initial output.

This indicates that neither interaction frequency nor time spent on the tool reliably predicts the degree of post-editing. The data point to a potential tendency toward acceptance of GPT output, which, in light of the nature of the text, is particularly worrying: the prevalence of zero-change submissions suggests that for many students, GPT functions less as a draft to be critically revised and more as a final product to be adopted with minimal intervention.

Translation of “act of God”

Students were asked to share their translations of specific translation problems in the post-task questionnaire. One of them was “act of God.” Their translations fall into three categories. The vast majority of students (13, 81.25%) opted for a literal calque, rendering the term as “acto(s) de Dios” (though with minor variation in number and capitalization: eight used the standard lowercase plural “actos de Dios”, two capitalized it as “Actos de Dios”, one used the singular “acto de Dios” and one wrote “actos de dios”, all lowercase). Only two students (12.5%) chose the conventional Spanish legal equivalent “fuerza mayor”: user2 rendered it as “casos de fuerza mayor” and user3 as “eventos de fuerza mayor”, while a single student (6.25%), user7, produced a descriptive paraphrase (“fenómenos naturales inevitables” [unavoidable natural phenomena]).

Looking back at the data, the three students who diverged from the literal calque show quite different interaction profiles. User3 (“eventos de fuerza mayor”) did ask specific terminology questions. This student explicitly directed GPT to *change* “actos de Dios” to “eventos de fuerza mayor” mid-session. This means the non-literal translation was not the result of GPT’s initiative or of a genuine exploratory question—the student already knew (or had searched it somewhere else) the functional equivalent and simply instructed the tool to apply it. User7 (“fenómenos naturales inevitables”) did not ask specific terminology questions, but asked GPT to *look deeper into the sentence “including but not limited to act of God”*—a meta-level reformulation request rather than a direct terminological query. This produced a naturalized, descriptive phrase, but not a legally grounded one, which suggests that open-ended prompts of this kind lead GPT towards paraphrase rather than towards established legal equivalence.

Self-assessment of legal translation competence, trust and risk awareness

The pre-task questionnaire indicated that the students had low confidence in their legal translation competence, with a mean score of 1.94 (n = 16) on a 1–5 scale. This is consistent with the profile of our undergraduate students with limited prior experience in legal translation. After the task, the mean score

modestly rose to 2.13 ($n = 16$). Only three students (18.75%) reported an increase in self-confidence and thirteen students (81.25%) remained at the same score. No student reported a decrease in self-confidence following the activity.

Similarly, the students reported a mean confidence in GPT's legal translation competence of 2.125 ($n = 16$). After completing the task, mean confidence in GPT rose to 2.75 ($n = 16$). The distribution shifted markedly toward the midpoint: ten students (63%) assigned a score of 3 ("medium confidence"). Three students scored 1, one scored 2, one scored 4 and one student (user15) assigned the maximum score of 5.

The written justifications that students provided for their post-task trust scores—both in relation to their own capacity and to ChatGPT—allowed us to identify three patterns.

The most repeated pattern was the difficulty in evaluating the quality of the output. Various students recognized that they did not have enough domain knowledge to know if GPT's translation was correct or not. For example, user4 wrote that "I think it has done an acceptable job, but I don't know to what extent it is a good translation or not" and user9 expressed something similar: "I still lack the knowledge to judge how it did it." These students were conscious of their limitations, but this consciousness did not help them to act in a different way.

A second pattern was related to the impact of AI use in the learning process. User11 reflected that the fact of delegating the translation to the tool had impeded real learning: "I did not actually work on the translation—it was done by the artificial intelligence. Therefore, I have learned absolutely nothing that could serve me in a future legal translation task." This is the unique case in which a student connected the use of AI directly to a loss in their own formation as a translator. In line with this reported no trust in GPT for translating legal texts and stated that, although ChatGPT had been useful for completing the exercise, they still had no confidence that an AI system could translate any text correctly or consistently without human intervention.

The third pattern was what we could call conditional trust, which is not put into practice. Some students said that the ChatGPT output was acceptable as long as it was revised by someone with knowledge, but these same students submitted the text without making any modifications. User15, for instance, wrote that the translation "could perfectly be used for a legal text, as long as there is post-translation editing", despite having changed zero words. This contradiction between what the student says and what the student does is especially relevant from a risk perspective, because it shows that having the correct procedural knowledge is not sufficient if the student lacks the competence to apply it.

Discussion

The findings reported in Section 4 (Results) can be understood on two levels. On the surface, they show some behavioral patterns in how undergraduate students of English used a generative AI tool during a timed legal translation task. On a deeper level, they relate directly to concerns that have influenced Pym's work on translator training, such as competence in a broader sense, the management of risk and uncertainty and how trust is adjusted in professional and technological contexts. The following discussion looks at each of these themes in relation to the research questions and considers what the methodological approach used in this study—based on API interaction logs instead of self-report—allows us to see in a new way.

Q1 asked what interaction and revision patterns can be observed in students' use of generative AI. The interaction data resist any linear interpretation that connects the number of interactions with the degree of critical engagement. The distribution of messages—polarized between a majority with two to four exchanges and two students with twenty-four each—might, at first glance, suggest that the more active users were also the more reflective ones. However, user2, who produced 24 messages and spent over 21 minutes on the tool, introduced only limited changes in the final text (7.4% of words modified), whereas user13 made the highest proportion of changes in the sample (29.1%) with just four interactions and almost no recorded time on the GPT interface. Similarly, several users with multiple interactions (e.g. users 4, 11 and 16) submitted the GPT translation without any modification. Thus, interaction frequency does not reliably indicate critical engagement, as both extensive and minimal use of the tool can result in either substantial revision or complete reliance on GPT output.

With regard to Q2 on how students formulate prompts when using generative AI for legal translation, our students formulated prompts in varied but generally simple ways when using generative AI for legal translation. Most rely on direct instructions, often requesting a full translation with minimal specification. A smaller group incorporates contextual information, such as the translation brief, tone, or target audience, resulting in more structured prompts. Some students engaged in iterative prompting, refining outputs through follow-up queries, including terminology questions or requests for reformulation. However, many prompts remained underspecified, with limited attention to legal nuances. Overall, prompting behavior ranged from instrumental, one-off requests to more exploratory and interactive strategies and reflected differing levels of engagement and competence. This echoes recent research showing that student translators' prompting practices tend to be intuitive and uninformed and that explicit AI literacy training is needed to foster more strategic use of GenAI tools in translation tasks (Zhang et al. 2025). It shows that students do not naturally use generative AI in a

strategically informed way, which might highlight a gap between access to the tool and prompting competence.

Regarding Q3, whether students' trust in AI-generated translations increased after the task, the pre- and post-task questionnaires showed that trust both in their own legal translation competence and generative AI competence rose after the task. This can be interpreted as a result of hands-on interaction with the tool, which may have fostered a sense of familiarity and perceived control, regardless of very little student-GPT interaction.

Building on interaction and text modification data patterns identified in Section 4, we identify four user profiles. The largest group consists of instrumental users, who submit a single request for full-text translation, receive the output and deliver it without further interaction or modification. In these cases, the tool is treated as a black-box translation machine and AI output is accepted as a final solution rather than a draft. A second group, referred to as exploratory users, engages in four to six exchanges, including occasional terminology queries or partial reformulations. However, this increased interaction does not result in substantial post-editing, as their engagement remains at the level of consultation rather than intervention. A third, less frequent profile is that of active post-editors, who interact iteratively with the system, question specific decisions and produce translations that differ significantly from the original output. Finally, one participant represents an autonomous user who uses the tool extensively as a parallel resource rather than a drafting aid, producing an independent translation. This profile aligns most closely with the pedagogical goals of critical AI use. These profiles should not be seen as fixed categories, but as context-dependent responses influenced by task difficulty and domain familiarity.

An observation that falls outside the main scope of this study but deserves attention is that no student selected Catalan as the target language, despite it being explicitly offered as an option. The reasons behind this unanimous choice cannot be determined from the available data, but several plausible explanations emerge. These include contextual factors, such as the instructions being provided in Spanish, individual language preferences and, more significantly, a potential lack of trust in GPT's ability to produce high-quality translations into a less represented language.

Perhaps the most significant finding emerges from the comparison of two data sources that, when considered independently, would suggest very different interpretations. The post-task questionnaires indicate that students are aware of the limitations of the tool, as several participants explicitly stated that the text produced by GPT would require expert revision before it could be used in a professional legal context. The interaction data, however, reveal that 50% of the participants submitted the translation provided by the system without any modification. The discrepancy between these two sets of data cannot be explained as an isolated mismatch. It constitutes a systematic

pattern across the sample and perhaps would not have been observable in a study based solely on self-reported data. The “act of God” example shows this competence–trust asymmetry clearly: most students accepted GPT’s translation (acto(s) de Dios or a paraphrase) and did not see that it was a serious legal error, even though they had said that such clauses should be checked by an expert. In studies on human-machine interaction, this phenomenon has been described as automation bias, understood as the tendency to follow the outputs of an automated system even when there are reasons to question them (Romeo and Conti 2025). In the context of this study, automation bias might be present not as blind trust in the system, since the students do not necessarily consider GPT to be infallible. Rather, it might take the form of epistemic surrender. Faced with the inability to verify whether the output is correct, the student implicitly treats it as acceptable. Zhang and Doherty (2025) argue that novice students may lack the competence to detect errors in AI translations, which can lead to over-reliance on these tools and hinder the development of critical thinking and creativity. However, students are likely to use generative AI for translation regardless. The key issue is therefore not whether to allow its use, but how to frame it pedagogically. Raising students’ awareness of their own limitations, as well as those of the tool, can foster more critical engagement and support the development of translator competence. This possibility opens up avenues for future research, for example, in an experiment consisting of various sessions, where students are allowed to see their results and act upon them.

The data reveals a paradox: far from generating additional caution, the specialized nature of the text appears to have had the opposite effect. Students did not trust GPT less because of the difficulty of the legal text, but rather more. Several participants explicitly acknowledged in the open-ended questions that they were aware that an error in a text of this nature could have serious legal consequences and that the result should be reviewed by an expert. And yet, 50% submitted the tool’s translation without any modification. This combination of simultaneous awareness of risk and inaction cannot be explained as negligence; it is the result of a specific competence asymmetry. As several students expressed, they did not feel capable of assessing whether GPT’s translation was correct or not. The specialized nature of the text, instead of increasing scrutiny, increased reliance on the tool precisely because it exposed the limits of the students’ own knowledge. In other words, the less you know, the more you trust, in this case, the machine (Tully et al. 2025). Trust does not stem from the belief that it performs well, but from the lack of any means to determine whether it is wrong.

Conclusions

A defining feature of our study lies in its reliance on behavioral rather than self-reported data. Instead of focusing on what students claim to have done, the analysis is grounded in what they actually did during their interaction with the tool. Access to interaction logs captured through the GPT API makes it possible to observe the translation process as it unfolds in real time, rather than reconstructing it retrospectively through questionnaires or interviews. This methodological orientation situates the study within the tradition of process-oriented research in translation and enables a more fine-grained understanding of student behavior, alongside studies with eye tracking, keylogging and think-aloud protocols.

The present study is subject to several limitations. First, it does not include a quality assessment of the final translations produced by the students. Such an analysis would have made it possible to determine to what extent the GPT API suggestions were incorporated into the final texts and whether their use had a measurable impact on translation quality. This limitation will be addressed in a separate article, in which the students' final translations will be analyzed in relation to their interaction data. A further limitation concerns the absence of full screen recordings. Such data would have made it possible to identify whether students relied on additional tools during the experiment, thereby offering a more comprehensive view of their workflow and enabling a more accurate interpretation of their interaction with the GPT system. Other limitations are the students' language and educational level. Taken together, these limitations point to the need for further research on interaction data in order to develop a more comprehensive account of how students engage with generative AI in translation tasks.

Taken together, the experiment shows that interaction logs uncover a clear gap between what students say about AI and what they actually do in legal translation tasks: many acknowledge the risks but still submit unedited GPT output. This suggests that GenAI-based tasks must be designed explicitly as learning opportunities that require students to scrutinize and revise AI translations, rather than simply use them as shortcuts. Further work with interaction data in other contexts will be needed to refine this account of how translation students learn to work with generative AI.

Declaration of AI use

In preparing this manuscript, the authors used Claude to reformulate and improve selected expressions. All AI-assisted content was subsequently reviewed, edited and verified by the authors, who take full responsibility for the accuracy and integrity of the final text.

References

- Almahasees, Zakaryia, Hussein Abu-Rayash, Samer Naser Olimat, and Dana Mahadin. 2024. "An analytical study on present perceptions of using ChatGPT in language translation". *Language Value* 17(2): 1-23. Universitat Jaume I ePress: Castelló, Spain.
<http://www.languagevalue.uji.es>
- Ayvazyan, Nune, Yu Hao and Anthony Pym. 2024. "Things to do in the translation class when technologies change. The case of AI text generation". In Yuhong Peng, Huihui Huang, Defeng Li (eds.) *New Advances in Translation Technology*. Springer.
- Belhassen, Saleh, and Achwak Hamda. 2025. "Translation Students' Reliance on and Trust in Artificial Intelligence for Successful Translation Projects: Opportunities, Challenges, and Implications". *Arab World English Journal for Translation & Literary Studies* 9 (2): 106-119.
<http://dx.doi.org/10.24093/awejtls/vol9no2.7>
- Bouyzourn, Kadija, and Alexandra Birch. 2025. "What Shapes User Trust in ChatGPT? A Mixed-Methods Study of User Attributes, Trust Dimensions, Task Context, and Societal Perceptions among University Students". *arXiv preprint*, <https://doi.org/10.48550/arXiv.2507.05046>
- EMT. 2022. European Master's in Translation Competence Framework. https://comission.europa.eu/system/files/2022-11/emt_competence_fw_k_2022_en.pdf. Visited March 2026.
- Muñoz Martín, Ricardo and Rojo López, Ana María. 2025. *Research methods in cognitive translation and interpreting studies*. Amsterdam: John Benjamins. <https://doi.org/10.1075/rmal.10.intro>
- Pym, Anthony. 2011. "Training translators". In Kirsten Malmkjær and Kevin Windle (Eds.) *The Oxford Handbook of Translation Studies*, pp. 475-489. Oxford: Oxford University Press.
- Pym, Anthony, Kirsten Malmkjaer, and Mar Gutiérrez. 2013. *Translation and Language Learning*. Luxembourg: European Commission.
- Pym, Anthony. 2015. "Translating as risk management". *Journal of Pragmatics*, 85, pp 67-80, ISSN 0378-2166.
<https://doi.org/10.1016/j.pragma.2015.06.010>.
- Pym, Anthony, and Nune Ayvazyan. 2017. "Linguistics, translation and interpreting in foreign-language teaching contexts". *The Routledge Handbook of Translation Studies and Linguistics*. 393-407.
 10.4324/9781315692845
- Pym, Anthony. 2025. *Risk Management in Translation*. Elements in Translation and Interpreting. Cambridge University Press.
- Pym, Anthony, and Yu Hao. 2025. *How to Augment Language Skills. Generative AI and Machine Translation in Language Learning and Translator Training*. Routledge.

- Romeo, Giuseppe and Conti, Daniela. 2025. "Exploring automation bias in human–AI collaboration: a review and implications for explainable AI". *AI & Soc* 41, 259–278. <https://doi.org/10.1007/s00146-025-02422-7>
- Su, Hanyu, Huilin Zhang, and Shihui Feng. 2026. "Comparing the Impact of Pedagogy-Informed Custom and General-Purpose GAI Chatbots on Students' Science Problem-Solving Processes and Performance Using Heterogeneous Interaction Network Analysis". <https://doi.org/10.48550/arXiv.2604.03022>
- Tully, Stephanie M.; Longoni, Chiara, and Appel, Gil. 2025. "Lower Artificial Intelligence Literacy Predicts Greater AI Receptivity". *Journal of Marketing*, 89 (5). <https://doi.org/10.1177/00222429251314491>.
- Zhang, Jia, and Stephen Doherty. 2025. "Investigating novice translation students' AI literacy in translation education". *The Interpreter and Translator Trainer*, 19(3–4), pp. 234–253. <https://doi.org/10.1080/1750399X.2025.2541478>
- Zhang, Jia, Xiaoyu Zhao, and Stephen Doherty. 2025. "Prompt engineering in translation: How do student translators leverage GenAI tools for translation tasks". *Proceedings of Machine Translation Summit XX 1*, pp. 420–431, Geneva, Switzerland. European Association for Machine Translation.

A constraints-first approach to CAT and generative AI in translator education: a simulated client project in a master's course

MAGGIE HUI

The Hong Kong Polytechnic University, Hong Kong

This paper presents a constraints-first, CAT- and GenAI-assisted simulated client project in a 13-week master's translation course (58 students in 12 groups). Students completed a group project requiring a consideration table, sequenced use of two CAT tools, capped trials (≤ 10 prompts) of multiple generative models and a four-pass MQM-lite QA loop (terminology \rightarrow accuracy \rightarrow fluency \rightarrow register/style). Analysis of one group's English-to-Chinese NVIDIA corporate-technical web-copy project documents typical GenAI risks through concrete artifacts and descriptive tallies. The study offers a replicable, standards-aligned classroom protocol supported by fully audited artifacts, while making no claims about efficiency gains or model superiority.

Keywords: Computer-assisted translation (CAT), generative AI, post-editing, translator education, simulated client project, terminology management, workflow design, strategic competence

Introduction

Professional translators today rarely work without tools (Christensen and Schjoldager, 2016). Standards such as ISO 17100 and ISO 18587, along with quality frameworks like Multidimensional Quality Metrics (MQM), now shape how translation is managed in companies. At the same time, translator education frameworks (EMT 2022 and PACTE) remind us that students need service provision, quality and strategic skills.

Yet there is still a noticeable gap. Industry expects auditable, constraint-driven workflows; many classrooms still offer only tool-familiarization exercises that lack real risk management. Recent research shows that the impressive fluency of neural systems can mask serious problems with accuracy and terminology (Koponen, 2016; Läubli, Sennrich, and Volk, 2018;

Yamada, 2019). These findings motivate staged quality checks that put terminology and accuracy first.

This article reports what happened when one master's-level student group ("Group 9") tackled a realistic English-to-Chinese (EN→CH) corporate-technical web-copy project. The source material consisted of 1,200 words from three closely related public NVIDIA blog posts announcing an "AI Blueprint for telco network configuration." Working under a demanding simulated client brief, the five students were required to strictly preserve all brand and product names (e.g., NVIDIA, NIM, AI Blueprint, BubbleRAN), consistently use a 17-item protected-term glossary they created themselves, maintain a concise corporate-technical register, avoid any unauthorized additions or summaries and provide full auditability of every AI-assisted decision.

The group employed three generative models (free ChatGPT, ChatGPT Plus and Microsoft Copilot) under a strict didactic cap of ≤ 10 prompts per model. They followed a carefully designed constraints-first workflow: an upfront consideration table to operationalize client requirements, sequenced use of two CAT tools (Trados Studio 2022 followed by Smartcat) for termbase and translation memory governance and a four-pass MQM-lite QA protocol (terminology → accuracy → fluency → register/style).

Drawing explicitly on Pym's (2013) emphasis on risk management as a foundational translator competence that outlasts any specific technology, the workflow prioritizes constraint identification, auditability and strategic decision-making over unchecked tool enthusiasm. In practice, this meant that typical GenAI risks, such as Copilot's tendency to insert speculative benefits not present in the source and the models' production of heavy pre-modifier structures that feel unnatural in formal Chinese, became visible and correctable at the appropriate stage.

Our aim is simply to describe what risks appeared and how the group handled them. We offer a ready-to-use classroom template, fully specified in Appendices A-C (the consideration table, the 17-item protected-term list, the three-stage prompting protocol and the MQM-lite revision log schema), that other programs can try and adapt.

The rest of this paper is organized as follows: It reviews the relevant literature, describes the research methods, presents the results from Group 9, discusses the implications, outlines the limitations and concludes.

Literature Review

Standards and quality frameworks shaping professional practice

Professional translation and post-editing are now guided by process and quality standards that spell out roles, documentation and verification steps.

ISO 17100:2015 sets out requirements for translation services, covering translator competence, project management, review processes (revision and proofreading) and the need for documented procedures (ISO, 2015). The standard separates responsibilities: the translator produces the first version, a reviser checks it against the source for accuracy and terminology, and a proofreader makes sure the target text reads correctly. This division of labor helps reduce individual bias and catch mistakes that one person might overlook.

ISO 18587:2017 focuses specifically on post-editing machine translation output. It distinguishes between light post-editing (mainly for quick understanding) and full post-editing (ready for publication) (ISO, 2017). The standard emphasizes meaning preservation, terminology consistency and proper quality control, particularly for full post-editing. It also reminds post-editors to watch out for typical machine-translation problems such as omissions, additions and misleading fluency. Together, these two standards offer a practical scaffold for designing classroom workflows: even when full commercial rigor is not possible, teachers can still adopt role separation, staged checks and audit trails.

For judging output quality, Multidimensional Quality Metrics (MQM) provides a category-based framework that makes quality issues visible and easy to classify (Lommel et al., 2014). MQM breaks quality into clear error types (for example, accuracy, terminology, fluency, style and consistency) and supports weighted scoring. In teaching situations, a simplified “MQM-lite” version (adapted from Lommel et al., 2014) that retains only the four most pedagogically relevant categories (terminology, accuracy, fluency and style/register) helps students focus on the most critical professional quality aspects while keeping the QA process manageable. The present study uses this reduced set to organize the revision logs.

Post-editing research: effort, error types and fluency-accuracy tensions

Research on post-editing (PE) has repeatedly shown that smooth surface fluency of machine translation (MT) output can hide real meaning errors. In this context, fluency refers to the linguistic quality of the target text (naturalness, grammatical correctness and readability), as opposed to accuracy, which concerns faithful preservation of the source meaning. Effort also tends to cluster around specific error types rather than simply the number of words. The three main types of post-editing effort (temporal, technical and cognitive) were first proposed by Krings (2001) and have since been widely adopted. Koponen’s (2016) review synthesizes findings from both statistical and neural MT systems, showing that cognitive effort does not always align with temporal or technical effort. A fluent but inaccurate sentence may require few keystrokes yet demand substantial mental effort to identify and correct.

Yamada's (2019) classroom study of students post-editing Google Neural Machine Translation found that although quality improved, students still struggled with typical NMT problems related to accuracy and terminology. They often trusted fluent output too readily and missed subtle meaning shifts or omissions. Yamada concluded that structured QA checklists and explicit training in error spotting are essential to overcome fluency bias. On a broader scale, Läubli et al. (2018) showed that human judges sometimes prefer more fluent MT output even when its adequacy is weaker, a tendency they called "fluency bias." When evaluators compared human and machine versions side by side, they frequently rated the machine text higher in fluency but lower in accuracy, yet their overall preference still leaned toward fluency.

In this context, accuracy refers to the faithful preservation of the source text's meaning (no unauthorized additions, omissions, or semantic shifts beyond approved translation strategies), aligning with ISO 18587's core requirement for meaning preservation and MQM's accuracy category (Lommel et al., 2014). These consistent findings explain why staged QA sequences (i.e., deliberate, ordered passes that prioritize terminology and accuracy before fluency or style) are so valuable: they force students to address the most critical risks first rather than being misled by fluent but inaccurate output. In the present study, Pass 1 (terminology) and Pass 2 (accuracy) are completed first, deliberately protecting the work from fluency-masked errors.

Translator-education perspectives: competence, risk and process transparency

Competence frameworks in translator education have long stressed that technical tool skills must sit inside larger service provision, quality and strategic competences. The EMT 2022 framework consists of six areas: language and culture, translation, technology, personal and interpersonal, service provision and thematic competence (EMT Expert Group, 2022). "Technology competence" goes beyond operating tools; it includes managing translation memories, termbases and machine translation systems, plus critically evaluating their output. "Service provision competence" covers project management, quality assurance and client communication. The present study puts both into practice: technology competence through dual-CAT sequencing and capped LLM trials and a simulated client brief captured in the consideration table, an upfront tool that operationalizes client constraints and register requirements

The PACTE model similarly highlights strategic competence as the central skill that coordinates tasks, including selecting procedures, spotting errors and solving problems (PACTE Group, 2017). Kiraly's (2000) social-constructivist approach shifted translator education toward authentic, collaborative problem-solving. The present study's simulated client project,

with its consideration table, role rotation and full artifact submission, directly embodies both PACTE's emphasis on strategic competence and Kiraly's principles by positioning students as active co-creators of knowledge in a realistic workflow.

Pym (2013) makes a strong case for training that puts risk management ahead of tool enthusiasm. Responding directly to assess risks (such as terminology drift, legal liability, or client dissatisfaction) and deciding where to spend effort is a lasting skill. In the present protocol operationalizes client constraints through the upfront consideration table and imposes a didactic prompt cap that forces students to make strategic decisions about when further prompting is no longer worthwhile.

Bowker and Buitrago-Ciro (2019) call for machine-translation literacy, which is the ability to understand what MT and similar technologies can and cannot do and how to use them responsibly in workflows. The present study extends this literacy to generative AI by adding prompt design, controllability checks and disciplined logging.

Taken as a whole, these three perspectives – Kiraly's emphasis on authentic collaborative tasks, Pym's focus on enduring risk management and strategic decision-making and Bowker and Buitrago-Ciro's call for critical technology literacy – provide a coherent pedagogical foundation for the constraints-first, auditable workflow examined in this study.

Emerging use of generative systems in translation workflows and education

Pym and Hao (2025) suggest viewing generative AI as an augmentation of human skill rather than a replacement. They recommend structured integration linked to clear learning outcomes: use AI for drafting and paraphrasing but always require human verification; teach students to document their AI use; and focus on tasks where AI reliably saves effort while avoiding high-risk tasks (such as legally binding texts) without close oversight.

Peer-reviewed studies on large language models (LLMs) in professional translation and translator education are growing, and several patterns have already emerged. First, LLMs produce fluent drafts very quickly, yet they are prone to hallucinations (factually wrong or invented content) and unauthorized additions (material not present in the source) (e.g., Guerreiro et al., 2023). Second, their ability to follow controlled terminology differs widely depending on the model and the prompt. Third, they show some sensitivity to register but can over-adjust or under-adjust if the prompt is not specific enough.

The existing standards (ISO 17100/18587) and quality taxonomies (MQM) were written before LLMs became widespread, but their process- and issue-based logic still works well. For instance, ISO 18587's requirement to "ensure that the meaning of the source text is preserved" directly tackles the unauthorized additions and omissions that are common LLM problems.

MQM’s “accuracy” category can flag hallucinated content, and its “terminology” category can catch drift. Educators therefore do not need to invent entirely new frameworks; they can adapt to the ones we already have.

Gap addressed by this study

Taken together, the literature offers consistent advice: classroom workflows should mirror professional process and quality standards, fluency alone should never be accepted as proof of quality and students need to develop strategic and risk-management skills alongside tool use. What is still missing, however, are concrete, replicable classroom protocols that combine CAT assets with generative systems in an auditable and standards-congruent way. Most existing studies focus either on CAT tools alone (e.g., Kornacki, 2018), on PE without LLMs (e.g., Koponen, 2016), or on LLM use without proper CAT governance (e.g., Guerreiro et al., 2023). The present single-case study addresses this gap by describing and analyzing one such orchestration in detail. It does not claim generalizable performance advantages over human-only or MT-only baselines; instead, it offers a practical template that other programs can adapt, evaluate and refine.

Methodology

Study design and scope

This is a single-case exploratory study focusing on one student group (“Group 9”) as they carried out a GenAI-supported CAT workflow translating from English into Chinese corporate-technical web copy taken from public NVIDIA blog material.

The purpose was descriptive: to document a constraints-aware classroom orchestration and to show how specific risks, such as unauthorized additions and terminology drift, became visible and were addressed. Claims about efficiency advantages over human-only or MT-only workflows, cross-group generalization and inferential statistics were deliberately kept out of scope. Group 9 was chosen because the group submitted complete artifacts as part of the course assessment and consented to in-depth analysis. It serves as a best-case illustration rather than a representative sample. Selection was based on completeness of artifacts and consent, which introduces a best-case bias in visible process quality and may under-represent typical error rates.

Participants and context

The study was conducted in a master’s-level translation course at the Hong Kong Polytechnic University in the 2025 autumn semester with 58 students in

12 groups and 39 contact hours over 13 weeks. Group members rotated roles across assignments. For this project, the five members of Group 9 collaborated jointly on terminology, drafting and QA tasks.

Source text and assets

The source text consisted of approximately 1,200 words from three closely related public NVIDIA web pages (one main blog post and two related GTC announcements) concerning the “AI Blueprint for telco network configuration” within the broader GTC-related communications context. These texts were publicly available on the official NVIDIA blog and contained only text (no images or other multimodal elements).

Group 9 created their own protected-term list and a bilingual glossary of 17 items and built the termbase and translation memory during the task. No external gold-standard term list was provided.

The consideration table containing client constraints and register information and the 17-item protected-term list used by Group 9 are reproduced in Appendices A and B, respectively.

Workflow orchestration (four-pass QA loop)

The project followed an instructor-designed constraints-first, four-pass QA loop adapted from the logic of ISO 17100/18587 and MQM:

- Pass 1 for terminology setup and enforcement: Build and verify the protected-term list and termbase; enforce glossary checks; repair all terminology issues before any drafting or stylistic work.
- Pass 2 for accuracy: Perform sentence-aligned source–target comparison to ensure meaning preservation (i.e., no unauthorized additions, omissions, or semantic shifts beyond approved strategies).
- Pass 3 for fluency and cohesion: Smooth sentence-level fluency and local coherence without altering protected terms or meaning.
- Pass 4 for register and style: Align output to the enterprise web register and style sheet; conduct final QA; prepare deliverables.

The group followed the three-stage prompting protocol described in Appendix C. Full prompt and revision logs were maintained throughout for auditability.

Tools and configuration

The group used two CAT environments: Trados Studio 2022 (desktop) for building the termbase and translation memory and Smartcat (cloud) for verifying consistency and terminology across the three related texts. Although the verification step could have been performed entirely in Trados Studio,

Smartcat’s cloud-based platform facilitated efficient cross-document consistency checking and real-time team collaboration. They first constructed the termbase and TM in Trados, then migrated the revised content to Smartcat to check term and TM reuse. Protected names and product lines were kept verbatim and glossary checks were used to flag any deviations. Numeric, date and tag consistency were handled through the standard CAT QA profiles.

A third-party MT plugin (Xiaoniu Translation) was used inside Trados due to version constraints; this is noted simply as a student-level detail rather than a general recommendation.

The exact settings applied by Group 9 are summarized in Table 1.

Table 1: CAT Tool Settings and QA Configuration Used by Group 9

Parameter	Setting
Desktop CAT	Trados Studio 2022
Cloud CAT	Smartcat (<i>exact version not recorded</i>)
TB/Glossary enforcement mode	Strict enforcement (manual override allowed)
Matching	Exact matching, case-sensitive for branded terms
Variants handled via	Manual addition to glossary; no regex used
Numbers and Measurements	Flag mismatches; unit consistency enforced for numerals
Tags/inline codes	Must match count and order (default QA profile)
Punctuation/spacing	Full-width for Chinese commas/periods; half-width for embedded English; non-breaking space before %
Terminology	Flag deviations from protected list; forbid paraphrasing of brand/product family names
Capitalization	Sentence-initial and proper-name rules; brand capitalization locked
Dates/numerics	Source pattern (e.g., “2024-03-25”) → target “2024年3月25日”
Repeated segments	Enforce consistency on repetitions; alert on divergent translations
Custom regex	None used in this project

Notes: All matches were reviewed manually; auto-acceptance was disabled. Match leverage (as reported by Smartcat): 101 % context matches: 12; 100 % exact matches: 8; repetitions: 23; fuzzy bands (95-99 %): 6; fuzzy bands (85-94 %): 11.

Generative model usage, data collection and analysis

Three models were compared: the free version of ChatGPT (powered by GPT-4o mini), ChatGPT Plus (powered by GPT-4o) and Microsoft Copilot (powered by a GPT-4-based model as of autumn 2025). The group applied the same three core prompts to each model: (1) baseline translation, (2) glossary-constrained revision and (3) consideration-table-constrained refinement. Copilot required three additional targeted prompts in this case.

The full three-stage prompting protocol and the additional prompts required by Copilot are provided in Appendix C.

A didactic prompt cap of ≤ 10 prompts per model was applied as a teaching constraint. Every prompt, model output and human edit was logged with timestamps.

Collected artifacts included the consideration table, protected-term list, TB/TMX exports, CAT QA notes, prompt log excerpts, segment-level before/after examples and the final bilingual deliverable.

Only descriptive tallies were produced (no composite scores). Categories were protected-term deviations, unauthorized additions, numeric/date mismatches, terminology inconsistencies (non-protected items) and fluency issues. Operational definitions were applied consistently in the revision log. The group logged observations; the instructor reviewed them for face validity. No inter-rater statistics were calculated because of the single-group, pedagogical nature of the study.

Ethics, replicability and deviations from ideal practice

The study was conducted under the ethical guidelines of The Hong Kong Polytechnic University for classroom-based pedagogical research. Students provided informed consent for the use of their anonymized artifacts in this analysis. Students were permitted to use GenAI for drafting and paraphrasing under strict constraints, but they remained fully responsible for meaning preservation and terminology compliance. Final submissions could include GenAI-originated text only if it passed Pass 1 (terminology) and Pass 2 (accuracy). Prompt and revision logs were submitted to the instructor as a mandatory requirement. External sharing of artifacts was limited to anonymized excerpts in accordance with institutional policy.

Templates and schemas for replication are provided in Appendices A-C. These include the consideration table (Appendix A), the protected-term list example (Appendix B) and the three-stage prompting protocol with the MQM-lite revision log schema (Appendix C). The main text also provides detailed descriptions of the CAT tool settings and annotated examples.

Deviations from ideal research practice were: (a) no human-only or MT-only baseline on the same text; (b) single-group design with no cross-group replication or inferential statistics; (c) no inter-rater reliability check for coding (all tallies are descriptive and pedagogical); and (d) time-on-task was estimated by the group (roughly 8-10 person-hours, excluding CAT training) but not measured systematically.

Results

Task description

Group 9 successfully completed the English-to-Chinese corporate-technical web-copy project. The simulated client brief was demanding and realistic: it required strict preservation of all brand and product names (such as NVIDIA, NIM, AI Blueprint and BubbleRAN), consistent use of the 17 approved Chinese terms, a concise corporate-technical register suitable for an official website, no unauthorized additions or summaries and complete auditability of every AI-assisted decision. The five group members collaborated jointly on terminology management, drafting and quality assurance. They estimated that the entire workflow took roughly 8-10 person-hours (excluding initial CAT tool training), although time-on-task was not measured systematically. As this was the students' first extended collaborative exercise combining CAT tools and generative AI, the project prioritized learning, tool familiarization, reflection and process discipline over translation speed; the study makes no claims about efficiency gains compared with human-only or MT-only translation.

Outcomes from the CAT tools

Even before any generative AI was introduced, the group's own termbase and glossary proved highly effective at catching basic inconsistencies. In Trados, the initial QA pass (conducted on human-translated segments during termbase and TM creation) identified several issues that would otherwise have carried through to later stages: mixed use of full-width and half-width semicolons in lists, inconsistent numeric formatting (for example, "2 million" versus "200万") and two accidental lower-case renderings of the branded product name "BubbleRAN." All of these were corrected at an early stage, demonstrating the value of building solid assets upfront.

The first text was used to construct the termbase and TM. After migrating the revised content to Smartcat, the group translated the remaining two texts from the same set of three. The platform reported 12 context matches (101 %), 8 exact matches (100 %), 23 repetitions and several fuzzy matches in the 85-99 % range. Segments containing protected terms were reused correctly and consistently across all three texts. No residual inconsistencies were reported. Although similar consistency could have been achieved in Trados alone, the dual-CAT approach (Trados for asset construction followed by Smartcat for cloud-based verification across texts) facilitated efficient cross-document checking and real-time team collaboration.

Generative AI performance under the capped prompt sequence

The group evaluated the three models (ChatGPT free version, ChatGPT Plus and Microsoft Copilot) using the identical three core prompts. This controlled prompting allowed direct comparison of model behavior under the same conditions.

ChatGPT Plus reached a quality level acceptable for post-editing (i.e., requiring only minor terminology and accuracy fixes before passing the four-pass QA loop) after only the third prompt and required no additional prompting. The model explicitly confirmed that it had received the glossary and even asked clarifying questions, such as whether “AI Blueprint” should always be rendered as “人工智能蓝图” even in headings. ChatGPT (free) also achieved acceptable quality after the third prompt, although its responses to multiple simultaneous constraints were slightly less precise than those of the Plus version. In contrast, Copilot still contained unauthorized additions and term drift after the third prompt, so the group issued three additional targeted prompts to bring it into line. These additional prompts specifically targeted the remaining unauthorized additions, terminology drift and style deviations that persisted after the core three-stage sequence (see Appendix C.2 and the concrete examples in the “Risk Profiles by Category” Section). The process stopped at six prompts for Copilot, well within the didactic ≤ 10 -prompt cap. Every prompt, model output and subsequent human edit was logged with timestamps and snapshots for full auditability.

Risk profiles by category

Terminology obedience

Comparison of the final model outputs against the 17-item glossary revealed clear differences in controllability. ChatGPT Plus showed the strongest adherence, replacing only 3 terms (17.6 % substitution rate) and producing zero unauthorized substitutions. ChatGPT (free) replaced 8 terms (47.1 %) but still recorded zero unauthorized substitutions. Copilot replaced 7 terms (41.2 %) and introduced 3 unauthorized substitutions, such as rendering “agentic AI” as “代理式人工智能” instead of the approved “智能体式人工智能”, “autonomous networks” as “自主网络” instead of “自治网络” and “incident ticket” as “事件工单” instead of “故障工单”. Some of the 17 glossary terms appeared more than once across the texts; thus the substitution rate refers to unique glossary items. All terminology deviations were corrected manually during Pass 1.

Example (Terminology – Copilot):

- Source: “...leveraging agentic AI capabilities...”

- Copilot output (after the core three prompts): “...利用代理式人工智能的能力...” (lit. “utilise **agent-style** artificial intelligence’s capability”)
- Approved glossary term: “智能体式人工智能” (lit. “**Agentic** AI”; the protected official NVIDIA term)
- Commentary: Copilot substituted a more common but unauthorized rendering. This deviation was caught and corrected in Pass 1 through strict glossary enforcement. It illustrates model-specific controllability challenges even under identical prompting conditions.

Unauthorized additions

The dedicated accuracy pass (Pass 2) identified six unauthorized additions across the three model outputs. Five originated from Copilot, including speculative benefits not present in the source and an invented subtitle “方案价值” (Solution Value). One minor addition came from ChatGPT (free). ChatGPT Plus produced none. All six additions were removed during Pass 2.

Example (Unauthorized Addition – primarily Copilot):

- Source: “The AI Blueprint provides a recipe for building autonomous networks.”
- Copilot output: “The AI Blueprint provides a recipe for building autonomous networks. This solution will accelerate the implementation of generative AI across telecom operations and reduce operational costs significantly.”
- Commentary: The added speculative benefits and cost-reduction claims were hallucinations not present in the source text. They were excised in Pass 2 through careful sentence-aligned accuracy checking. This example clearly demonstrates why a dedicated accuracy pass must precede any fluency or style editing.

Mistranslations and modifier scope errors

Both Copilot and ChatGPT (free) occasionally mishandled modifier scope. In one sentence describing announcements at NVIDIA GTC Paris, the models incorrectly attached the relative clause “showcasing...” to the conference location rather than to the announcements themselves. The group corrected this manually in Pass 2. A parallel terminology error occurred with “incident ticket”, where both Copilot and ChatGPT Plus initially used the wrong glossary form.

Fluency and register issues

Even after the dedicated fluency and register passes, all three models required substantial human restructuring to achieve natural, enterprise-level Chinese.

The dominant problem was heavy pre-modification. This refers to long strings of attributes placed before the subject, which feels awkward in formal Chinese web copy.

Example (Heavy Pre-modification for all models, EN→CH specific):

- Source: “Norway-based Telenor Group, which serves over 200 million customers globally, is the first telco to integrate the AI Blueprint...”
- Model outputs (all three): “总部位于挪威、服务全球超过2亿客户的Telenor集团是首个将人工智能蓝图集成到电信网络配置中的电信运营商.....”
- Human revision (Passes 3–4): “Telenor 集团总部位于挪威，全球服务客户超过两亿。该集团是首个集成人工智能蓝图的电信运营商.....”
- Commentary: The models faithfully reproduced the English-style complex subject with stacked pre-modifiers, resulting in a sentence that feels unnatural and overloaded in Chinese. Breaking it into shorter, clearer clauses significantly improved readability and register appropriateness. This structural challenge is particularly notable in EN→CH post-editing and further justifies placing terminology and accuracy checks before fluency and register passes.

Quantitative summary of post-editing effort

Analysis of Group 9’s revision logs (as reviewed by the instructor) recorded a total of 78 edits across the three texts. The distribution by MQM-lite category was as follows:

Table 2: MQM-lite Edit Distribution

MQM-lite Category	Number of Edits	Percentage of Total
Terminology (protected-term deviations)	24	31%
Accuracy (unauthorized additions/omissions)	11	14%
Fluency (ungrammatical or non-idiomatic)	28	36%
Style/Register (pre-modifications, etc.)	15	19%
Total	78	100%

Notes: Fluency edits were the most numerous, followed by terminology edits. Accuracy edits totaled 11 (6 unauthorized additions and 5 omission or modifier-scope errors). These accuracy issues originated only from Copilot, while ChatGPT Plus produced none.

Discussion

Visibility of typical risks under a staged workflow

The Group 9 case shows that a constraints-first, four-pass workflow (terminology → accuracy → fluency → register/style) makes typical GenAI risks visible and manageable at an early stage. First, protected-term deviations were caught in Pass 1 because the glossary was enforced before any stylistic polishing began. Second, unauthorized additions, a common problem with generative models especially Copilot, were identified in Pass 2 through careful sentence-aligned accuracy checks. Third, numeric/date mismatches and modifier scope errors (for example, incorrectly attaching the relative clause “showcasing...” to “GTC Paris” instead of to the announcements) were flagged and corrected in the same pass. Without this deliberate staging, a substantial number of these errors would have survived into the final deliverable, hidden behind otherwise fluent output. This finding supports Yamada’s (2019) observation that NMT’s surface fluency can mask adequacy problems and suggests it extends the same insight to LLM-assisted drafting.

Controllability and prompt sequencing

Across the three models, the sequence of prompts proved far more important than the precise wording of any individual prompt. When the group first tried refining the style before locking down terminology, the models, especially Copilot, frequently reintroduced term drift (for example, changing “人工智能蓝图” back to “AI Blueprint”). Once the sequence was reordered to place terminology and accuracy first, the amount of drift decreased markedly. This finding highlights a clear pedagogical principle: terminology and meaning should be treated as non-negotiable invariants, while fluency and register can be adjusted as later, more flexible layers. For instructors, this principle is far more practical and transferable than relying on model-specific prompting techniques.

Prompt cap as a teaching constraint, not a performance metric

The ≤10-prompt cap was introduced purely as a teaching tool. It forced students to budget their interactions carefully and to decide which constraints were worth enforcing. In practice, ChatGPT Plus reached an “acceptable for post-editing” state within three prompts, while Copilot needed six. The cap prevented endless trial-and-error tweaking and encouraged students to move on to manual editing when a model proved uncooperative. This mirrors real-world professional practice, in which translators routinely decide when further machine assistance is no longer cost-effective and must shift to post-editing (Pym 2013). Importantly, the cap is not presented as a claim about model

quality; it is simply a classroom device to build strategic competence (PACTE Group, 2017) and risk-management skills (Pym, 2013).

Alignment with professional process logic

The workflow closely echoes the role separation described in ISO 17100 (translator → reviser → editor) and ISO 18587's strong emphasis on meaning preservation. However, the classroom setting differs from commercial practice in several important ways: grading incentives, fixed deadlines and the lack of formal client-vendor contracts. For this reason, we do not claim that the student workflow is equivalent to vendor-grade practice. Nevertheless, the underlying logic of staging quality checks and maintaining full audit trails through prompt logs and revision logs is highly transferable. Students learned to justify every significant change by explicitly linking it to an MQM-lite category (e.g., "this was corrected under Accuracy because it introduced an unauthorized addition"), a skill that is directly relevant to professional quality management and audit requirements (Lommel et al., 2014).

Pedagogical hypotheses for future testing

The present single-case study suggests several pedagogical hypotheses that could be examined in future replications with larger samples:

- H1 (Sequencing effect): In EN→CH corporate-technical translation, applying terminology and accuracy gates before style-related prompts reduces the reintroduction of term and meaning errors compared with a style-first sequence.
- H2 (Asset effect): Groups that construct a minimal protected-term list and termbase upfront produce fewer protected-term deviations than groups that do not.
- H3 (Audit effect): Requiring prompt and revision logs improves students' ability to correctly classify their own edits into MQM-lite categories (accuracy, terminology, fluency, style) by the end of the semester.

Recommendations for instructors

This case yields four practical recommendations for instructors:

- (1) Front-load constraints. Require students to complete a consideration table before any drafting or prompting and treat it as the single source of truth for all QA decisions. This simple step forces students to articulate client needs and register expectations explicitly, making subsequent tool and AI use far more purposeful.
- (2) Sequence, do not just prompt. Enforce the order: terminology → accuracy → fluency → register/style. Allow style-related prompts

only after protected terms have been locked. In practice, this dramatically reduces the reintroduction of term drift and unauthorized additions.

- (3) Cap and compare. Use a modest prompt budget (e.g., ≤ 10 per model) so students can discover differences in model controllability and learn to value systems that respect constraints rather than endlessly tweaking prompts.
- (4) Assess process, not just product. Base grading on artifacts such as prompt logs, revision logs tagged with MQM-lite categories and TB/TMX files. This rewards disciplined workflows and strategic decision-making rather than final polish alone and mirrors professional audit requirements.

We make no claims that any particular model is universally superior, that the observed error counts would replicate in other texts or language pairs, or that this workflow is more efficient than human-only translation. These boundaries are essential for interpreting the case correctly.

Limitations

The single-case study, while providing a detailed illustration of the protocol in action, has several limitations. The genre (corporate-technical web copy) and language direction (EN→CH) may limit generalizability; error profiles could differ substantially in legal, medical, literary, or other language pairs. The classroom context, with its time pressures and pedagogical focus, differs noticeably from commercial practice. The outcomes should therefore be viewed as pedagogical observations rather than vendor-grade evidence.

Descriptive tallies rely on the group's self-reported revision logs, reviewed by the instructor for face validity only, with no inter-rater reliability assessment. Model behavior is time-sensitive due to provider updates and stochastic sampling. Full artifacts cannot be publicly released due to institutional policies, limiting replication to the provided templates, settings and anonymized excerpts.

These constraints mean the study documents visible risks and workflow manageability in one pedagogical instance but does not support broader quantitative inferences or cross-context claims. The methodological choices underlying these constraints, such as the single-group design, absence of baselines and lack of inter-rater reliability, were already explained and justified in the “Ethics, Replicability and Deviations from Ideal Practice” Section.

Conclusions

This single case demonstrates that a constraints-first workflow combining CAT assets with a fixed prompt sequence makes typical risks visible at an early stage and correctable before stylistic work begins. These risks include protected-term deviations, unauthorized additions and numeric mismatches. The glossary-first setup and dedicated accuracy pass played a central role in surfacing and repairing these issues, while achieving register-appropriate Chinese still required substantial human restructuring. The main practical contribution is a replicable classroom protocol supported by concrete artifacts (consideration table, TB/TMX files, prompt and revision logs and CAT-QA settings) that other programs can readily adapt and test.

Further research should strengthen the evidence base through baseline comparisons, rigorous inter-rater reliability assessment on stratified samples and replication across different groups, genres and language directions using identical CAT-QA settings.

Declaration of AI use

In preparing this manuscript, the author used Grok for language editing, literature search and checking referencing format. All AI-assisted content was reviewed, edited and verified by the author, who takes full responsibility for the accuracy and integrity of the final text.

References

- Bowker, L. and J. Buitrago-Ciro. 2019. Machine translation and global research: Towards improved machine translation literacy in the scholarly community. Emerald Publishing.
- Christensen, T. P. and Schjoldager, A. 2016. "Computer-aided translation tools – the uptake and use by Danish translation service providers". *The Journal of Specialised Translation*, 25, 89-105.
- EMT Expert Group. 2022. European Master's in Translation: EMT competence framework 2022. European Commission.
- Guerreiro, N. M., Alves, D., Voita, E., and Martins, A. F. T. 2023. "Hallucinations in large multilingual translation models". *Transactions of the Association for Computational Linguistics*, 11, 1500-1517.
- ISO. 2015. ISO 17100:2015 – Translation services: Requirements for translation services. International Organization for Standardization.
- ISO. 2017. ISO 18587:2017 – Translation services: Post-editing of machine translation output – Requirements. International Organization for Standardization.

- Kiraly, D. 2000. *A social constructivist approach to translator education: Empowerment from theory to practice*. St. Jerome.
- Koponen, M. 2016. "Is machine translation post-editing worth the effort? A survey of research into post-editing and effort". *The Journal of Specialised Translation* 25: 131-148.
- Kornacki, M. 2018. *Computer-Assisted Translation (CAT) Tools in the Translator Training Process*. Peter Lang.
- Krings, H. P. 2001. *Repairing texts: Empirical investigations of machine translation post-editing processes*. Kent State University Press.
- Läubli, S., R. Sennrich, and M. Volk. 2018. "Has machine translation achieved human parity? A case for adequate evaluation". In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 4791-4796. Association for Computational Linguistics.
- Lommel, A., Uszkoreit, H., and Burchardt, A. 2014. "Multidimensional Quality Metrics (MQM): A framework for declaring and describing translation quality metrics". *Revista Tradumàtica: Tecnologies de la Traducció*, 12, 455–463.
- PACTE Group. 2017. "PACTE translation competence model: A holistic, dynamic model of translation competence". In A. Hurtado Albir (ed.) *Researching translation competence by PACTE Group*, 35-42. John Benjamins.
- Pym, A. 2013. "Translation skill-sets in a machine-translation age". *Meta*, 58(3), 487–503.
- Pym, A. and Y. Hao. 2025. *How to augment language skills: Generative AI and machine translation in language learning and translator training*. Routledge.
- Yamada, M. 2019. "The impact of Google Neural Machine Translation on post-editing by student translators". *The Journal of Specialised Translation* 31: 87-106.

Appendix A. Group report: consideration table

Consideration	Remarks
Genre	Technical informative text: website
Orientation	Information-centered
Intended Text Function (main)	Informative*** Vocative** (hidden)
User-friendliness	***
Intended Readership (geographical)	People from China
Intended Readership (others)	Tech enthusiasts and potential customers
Register: Field	Technical texts
Register: Mode	Written and published on the NVIDIA official website
Register: Tenor	Tech giant and its users
Overall Register	Neutral-to-Formal
Translation Purpose	Promotion of generative AI in NVIDIA Support NVIDIA's branding and communication goals in the local market
Translation Approach	Communicative Translation
Translation Strategy	Paraphrase, explicitation, generalization, limited omission/addition (elaboration) and adjustment (e.g., sentence order) Except for literalism

Appendix B. Group report: protected-term list example

This is the exact protected-term list used by Group 9, which was created in Excel, imported into Trados and Smartcat and enforced strictly in all AI prompts.

No	English Term	Chinese Term	Notes / Protection Rule
1	AI Blueprint	人工智能蓝图	Protected brand term: never leave in English
2	AI	人工智能	Replace all standalone occurrences
3	NIM microservices	NIM 微服务	Keep “NIM” in English
4	Telco network configuration	电信网络配置	Core technical phrase
5	Agentic AI	智能体式人工智能	Official NVIDIA term
6	GTC Paris	GTC 巴黎	Event name: keep English acronym
7	5G O-RAN	5G O-RAN	Keep English acronym
8	Autonomous network	自治网络	Preferred translation
9	NVIDIA AI Enterprise	NVIDIA AI Enterprise	Brand name: no translation
10	AI Refinery	AI Refinery	Platform name: keep English
11	BubbleRAN	BubbleRAN	Partner platform: keep English
12	Telenor Group	Telenor 集团	Company name: keep English
13	Accenture	埃森哲	First occurrence only
14	Incident tickets	故障工单	Standard industry term
15	Deployment tools	部署工具	Consistent technical term
16	Accelerated computing	加速计算	Core NVIDIA technology term
17	Generative AI	生成式人工智能	Official Chinese equivalent

Appendix C. Group report: prompt design and methodology

C.1 Core three-stage prompting protocol

The group applied the same controlled three-stage prompting sequence to all three models using the versions accessible to end-users at the time of data collection (September-December 2025): ChatGPT (free tier) powered by GPT-4o mini, ChatGPT Plus powered by GPT-4o and Microsoft Copilot powered by the GPT-4-based model then current.

The sequence was designed to mirror real-world professional constraints:

- (1) Baseline translation: raw source text with minimal instruction.
- (2) Terminology-constrained revision: revision using the bilingual glossary created in the CAT-tool stage.
- (3) Consideration-table refinement: full alignment with the client Consideration Table (terminology accuracy, accuracy, corporate-technical register, logical clarity and conciseness).

The exact prompts used in each stage are shown in Table C1.

Table C1: Core three-stage prompting protocol

Stage	Prompt	Purpose
1. Baseline	Translate the following text into Chinese.	Obtain each model's default, unguided output.
2. Terminology Control	Please revise the above translation according to the provided bilingual glossary.	Test adherence to domain-specific terminology developed in Trados/Smartcat.
3. Consideration-Table Refinement	Please refine and improve the translation according to the Consideration Table. Ensure terminology accuracy, corporate-technical register, logical clarity and conciseness.	Align output with all client constraints and professional standards.

Notes: These three core prompts were applied identically to all models. Copilot required three additional targeted prompts to resolve persistent deviations (see C.2 below).

C.2 Additional targeted prompts (Copilot only)

Because Copilot continued to deviate after the core prompts, the group issued three supplementary prompts:

- (4) Please adjust the style further to match the language style of NVIDIA’s official Chinese website.
- (5) When making revisions, do not deviate from the glossary and Consideration Table I provided earlier. Please verify and correct accordingly.
- (6) Please change “AI Blueprint” to “人工智能蓝图” and change “AI” to “人工智能”.

C.3 Revision log schema (MQM-lite categories)

All human edits were recorded in a standardized revision log using the following MQM-lite categories. Each edit was tagged with the model that produced the original error and the pass in which it was corrected.

Category	Definition	Example
Terminology	Deviation from the protected-term list or glossary (including unauthorized substitution or drift)	“AI Blueprint” rendered as “AI 蓝图” instead of the required “人工智能蓝图”
Accuracy	Unauthorized addition, omission, or scope error that changes meaning	Addition of “加速生成式人工智能的落地应用” not present in the source text
Fluency	Ungrammatical, awkward, or non-idiomatic Chinese	Nested parentheses or pre-modifier stacking that violates natural Chinese word order
Style/Register	Failure to match the required corporate-technical register of NVIDIA’s official Chinese website	Overly colloquial phrasing or inconsistent formality

Usage instructions for replicability

- Log every prompt, model output and human edit with timestamps.
- Tag each edit with the MQM-lite category above.
- Record which model introduced the error, and in which pass it was fixed.
- Export the log as an Excel or CSV file together with the final TMX/TB for full auditability.

Beyond perceptual failure: cognitive load and number renditions in technology-mediated English–Turkish simultaneous interpreting

ENES EKİCİ and ŞEYDA ERASLAN

Dokuz Eylül University, Türkiye

This article investigates whether real-time simulated ASR visual support improves numerical rendition accuracy in English-Turkish simultaneous interpreting among student interpreters. Using a within-subjects experimental design, participants interpreted number-dense source texts with and without ASR-simulated captions. The findings are analyzed through Gile’s Effort Model and Seeber’s Cognitive Load Model. Results indicate that visual support did not eliminate numerical errors but shifted the error profile, reducing omissions while increasing syntactical errors – suggesting a redistribution of cognitive resources rather than an overall reduction in processing load.

Keywords: simultaneous interpreting, numerical rendition, computer-assisted interpreting, Cognitive Load Model, English–Turkish language pair, interpreter training

Introduction

Numbers have long been recognized as one of the most well-documented (Braun and Clarici 1996; Mazza 2001; Pinochi 2009; Mead 2015) sources of failure in simultaneous interpreting. Unlike lexical items, they resist prediction, offer little contextual redundancy and demand a kind of cognitive processing that sits uneasily alongside the already considerable demands of real-time interpreting. Gile’s (1995) Effort Model frames this well: when Listening, Memory and Production efforts collectively exceed available capacity, something has to be sacrificed and numbers are frequently what gets sacrificed.

One response to this problem has been technological. The broader “technological turn” in interpreting (Fantinuoli 2018) has produced a

generation of Computer-Assisted Interpreting (CAI) tools that move beyond preparation aids and glossary management to intervene directly during the interpreting task. The most recent of these tools use automatic speech recognition (ASR) to flag precisely the kind of high-density triggers, such as numbers, proper names, technical terms, that tend to push interpreters past their processing limits. Early experimental results are encouraging: studies have reported number rendition accuracy rising from under 60% to over 85% when real-time visual support is provided (Desmet et al. 2018; Defrancq and Fantinuoli 2021). However, most of this work has been conducted with professional interpreters or in language pairs other than English–Turkish, leaving open the question of how student interpreters respond to this kind of support.

The question this study sets out to answer is a practical one: does providing student interpreters with a visual display of numerical information actually help the rendition of numbers? And if so, how does it help—does the gain come simply from having the number visible, or does it reflect something deeper about how the brain handles multimodal input? Seeber’s (2011) Cognitive Load Model offers a theoretical basis for optimism here. If visual-spatial and auditory-verbal channels operate as relatively independent resource pools, then a well-designed visual aid should relieve rather than compete with the auditory processing demands of interpreting. Whether that theoretical prediction holds in practice, specifically in the English–Turkish language pair and among student interpreters, is what this study examines.

English–Turkish interpreting presents a particularly interesting test case. The structural distance between the two languages means that even a correctly perceived figure can be incorrectly produced in the target language and this can happen in more than one way. Some figures map directly across the pair—English “4.2 trillion” corresponds to Turkish “4,2 trilyon”, preserving both magnitude and structure—while others require structural reformulation: English reads years in paired two-digit units (“nineteen eighty-five”), whereas Turkish renders the full cardinal value (“*bin dokuz yüz seksen beş*”), demanding real-time decomposition rather than transcoding. A further difficulty concerns the order of magnitude rather than the digits themselves. English “million” and “billion” correspond to Turkish “*milyon*” and “*milyar*”, near-identical forms that compete under time pressure, so that selecting the wrong scale word leaves the digits intact but relocates the entire figure across magnitudes. Although such a slip resembles a simple lexical substitution on the surface, the failure is structural—the number is reconstructed on the wrong scale. A figure can therefore be heard and understood correctly yet still fail at the production stage, for reasons that have nothing to do with perception. This distinction between perceptual failure and structural failure matters a great deal for how we understand what challenges in interpreting technology can and cannot address.

To investigate this, six senior interpreting students from the Department of Translation and Interpreting at a state university in Türkiye performed simultaneous interpretation of two numerically dense English speeches: one without any technological support and one accompanied by a simulated ASR display showing numbers in real time. The comparison allows for a direct assessment of whether visual support improves rendition accuracy and—equally importantly—whether it changes the type of errors students make.

By framing the current study within the broader context of the “technological turn” (Fantinuoli 2018), this research addresses the critical need for empirical data on how process-oriented tools affect the interpreter’s cognitive performance in the English-Turkish language pair. While traditional strategies for managing numbers, such as the assistance of a boothmate or manual note-taking, have long been the standard, the integration of real-time visual support via Computer-Assisted Interpreting (CAI) represents a fundamental shift in the “cognitive resource footprint” of the task (Seeber 2017; Mazza 2001). Before detailing the experimental findings of the English–Turkish rendition, the following sections review the literature and theoretical frameworks that prop up this transition. This includes a historical analysis of how technology made simultaneous interpreting the new norm, the evolution of CAI generations and the cognitive models, moving from Gile’s (1995) bottleneck theory to Seeber’s (2011) multimodal approach.

Literature review and theoretical background

Technology and the Technological Turn in interpreting

Simultaneous interpreting as a professional practice is inseparable from the technology that made it possible. The booth, the headset, the relay system—these are not neutral tools but the conditions under which the profession came into existence. The Filene-Finlay system of the 1920s and its adaptation at the Nuremberg Trials (1945–1946) established that real-time multilingual communication was not only possible but scalable (Baigorri-Jalón 2014; Pöchhacker 2015). When the United Nations formally adopted simultaneous interpreting as a permanent service through Resolution 152 (II) in 1947, it did so for pragmatic reasons—time and efficiency—not because the profession had resolved its theoretical or cognitive questions (United Nations General Assembly 1947; Pöchhacker 2015).

What has changed in recent decades is the locus of technological intervention. Early technologies shaped the setting of interpreting: the booth, the transmission hardware, the acoustic environment. The current wave, which Fantinuoli (2018) terms the “technological turn”, reaches further inward, targeting the cognitive sub-processes of the task itself. CAI tools now attempt to assist interpreters not before or after an assignment but during it, in

real time. This is a qualitatively different kind of intervention, and it demands a correspondingly different theoretical framework for evaluation.

CAI tools and interpreter support during the task

The evolution of CAI tools is typically divided into three generations and the distinction between them is not merely technical but conceptual. First-generation tools—electronic glossaries like Interplex or Terminus—were essentially digitized reference materials, useful for preparation but passive during the task itself (Fantinuoli 2018). Second-generation tools introduced corpus-based preparation and smarter search functions, still oriented primarily toward the pre-task phase. The third generation is where the shift becomes genuinely significant: ASR-driven tools that operate in real time, surfacing high-density triggers, such as numbers, proper names and technical terms, at the moment they occur in the source speech (Fantinuoli 2018; Lu and Fantinuoli 2025).

The empirical case for these tools, at least where numbers are concerned, is difficult to dismiss. Desmet et al. (2018) reported number rendition accuracy rising from 56.5% to 86.5% with real-time visual support. Defrancq and Fantinuoli (2021) found similar gains with InterpretBank, with accuracy moving from 67.7% to 90.2% alongside a significant reduction in omissions. These are large effects by any standard and they suggest that the perceptual problem, failing to capture a number in the first place, is genuinely addressable through technology. What remains less clear and what this study attempts to prove, is whether solving the perceptual problem is sufficient, or whether structural and production-level failures persist even when the number is visible.

Numbers in simultaneous interpreting: Errors and strategies

The difficulty of numbers in simultaneous interpreting is well-documented and, at this point, largely uncontested. What is worth examining more carefully is why they are difficult, because the answer shapes what kind of support is actually useful.

Numbers are not simply unfamiliar words. They carry high information density with almost no contextual redundancy (Mazza 2001)—if you miss “forty-seven million”, there is no surrounding syntax that helps you reconstruct it. They also engage neural pathways that partially dissociate from those involved in ordinary language processing: exact numerical computation recruits both verbal memory networks and bilateral parietal areas associated with magnitude representation (Dehaene et al. 1999), a dual demand that has no real equivalent in lexical processing. Under Gile’s (1995) Effort Model, this cost is straightforward to describe: numbers intensify the Listening and

Short-term Memory efforts at precisely the moments when capacity is most constrained, pushing the system toward saturation.

Error rates reflect this. Braun and Clarici (1996) documented mean error rates approaching 70% among student interpreters. Professional interpreters fare better but remain far from reliable: Timarová (2012) reported error rates of approximately 40%, a figure consistent with the 43% observed at higher delivery rates by Desmet et al. (2018). The dominant error type in the unsupported conditions is omission—the interpreter drops the number rather than risk a more visible failure. In Pym’s (2025) terms, this is a risk-management decision: the interpreter forgoes a high-risk item rather than produce a conspicuously erroneous figure. This is also consistent with Gile’s framework: omission preserves the flow of the target speech at the cost of one piece of information. Approximation serves a similar function, rounding “47” to “around 50” to reduce local cognitive load while maintaining communicative continuity (Pinochi 2009).

The traditional human solution, a boothmate who notes figures as they appear, effectively externalizes the Memory Effort, offloading it to a visual aid. The importance of this collaboration was underscored by Arzık Erzurumlu and Demir (2022), who found that numerical accuracy declined sharply in Turkish broadcasts of the 2020 American Presidential Debates when COVID-19 social distancing requirements prevented interpreters from assisting each other in the booth. CAI tools are, in a sense, a technological formalization of this same principle—and one that does not depend on the physical proximity of a colleague.

Multimodality, ASR and cognitive load

The dominant theoretical framework for cognitive load in interpreting has long been Gile’s Effort Model, and it remains indispensable as a descriptive tool. But it has a significant limitation: it treats attention as a single, undifferentiated resource. Under this view, any additional input—including a visual display—competes with the existing demands of listening and production and is therefore a potential source of overload rather than relief (Gile 2009).

This is where the model is demonstrably insufficient. The assumption of a unitary resource pool does not sit well with what we know about how the brain handles multimodal input. Wickens’ (2002) Multiple Resource Theory proposes that the human information-processing system contains several relatively independent resource pools, organized by modality, processing code and stage. Two tasks interfere significantly only when they draw from the same pool; tasks that engage different modalities can, in principle, be performed concurrently at a much less cost.

Seeber (2011) builds directly on this foundation to offer a Cognitive Load Model specific to simultaneous interpreting. His “conflict matrix”

makes the key point clearly: visual input is not inherently a competitor for the auditory-verbal resources that interpreting primarily demands. If a visual signal is synchronous and congruent with the auditory stream, as a live number display would be, it can function as cognitive support rather than additional load. This reframes the question entirely. The issue is not whether to introduce visual input, but how to design it so that it operates in a complementary rather than competing register.

Empirical support for this position is growing. Eye-tracking data suggest that interpreters strategically coordinate visual attention with auditory input during text-supported SI: professional interpreters show an ear-lead-eye pattern that enables efficient multimodal synchronization (Seeber et al. 2020), while trainee interpreters have been found to actively seek out numbers and proper names on screen when live captioning is available (Yuan and Wang 2023). ASR-generated captions have been shown to reduce cognitive load specifically for the kind of triggers, such as numbers and specialized terminology, where auditory processing is most vulnerable (Defrancq and Fantinuoli 2021). Taken together, this evidence supports Seeber's model over Gile's on this particular question: visual support, when well-designed, redistributes cognitive demand rather than compounding it.

Research questions

Guided by the theoretical frameworks of Gile (1995) and Seeber (2011) and the historical transition toward technology-mediated interpreting, this study addresses the following research questions:

RQ 1: What is the level of numerical rendition accuracy among student interpreters in the English–Turkish language pair?

RQ 2: Which strategies (e.g., omission, approximation, transcoding and delaying) do students employ when interpreting numbers?

RQ 3: Does the provision of technological support (simulated ASR) significantly affect the accuracy of numerical rendition?

Regarding RQ3, existing literature on multimodal facilitation suggests that technological support enhances performance by allowing for a “cognitive load transfer” between modalities. Seeber (2012) notes that interpreters actively seek visual stimuli to complement auditory speech, particularly when processing numerals. Experimental data from professional cohorts show that accuracy rises when supported by visual aids. Eye-tracking evidence further confirms that numerical processing imposes a disproportionately high cognitive load compared to surrounding content, as reflected in significantly longer fixation durations on numbers than on the items they refer to (Korpál and Stachowiak-Szymczak 2018). In one pilot study with trainees, number accuracy increased from 56.5% to 86.5% when supported by real-time visual displays (Desmet et al. 2018).

Methodology

This study employs a within-subjects experimental design to examine how real-time visual support affects the rendition of numerical information in English–Turkish simultaneous interpreting. Each participant served as their own control, interpreting one speech without technological assistance and one with simulated ASR support. This design was chosen to minimize the influence of individual variation in baseline interpreting ability, a particularly important consideration given the small sample size.

Participants

Six senior students ($N = 6$) from the Department of Translation and Interpreting at a major Turkish state university participated in the study. The experiment was conducted at the end of their eighth and final semester, ensuring that all participants had completed the full simultaneous interpreting curriculum, which includes Simultaneous Interpreting I and II in fall and spring terms respectively, each meeting 3 hours per week. All reported Turkish as their native language and English as their primary working language. The sample size is admittedly modest, and the findings should be read as exploratory evidence that warrants replication with a larger cohort rather than as definitive claims about the population of English–Turkish student interpreters.

Materials and stimuli

Two English-language speeches were prepared as experimental stimuli. The speeches were scripted to reproduce the register, structure and lexical density typical of conference presentations on the chosen topics, and were recorded by one of the researchers, a speaker with professional experience using English as a working language. Speech rate was controlled at 86 words per minute across both conditions. The target rate was achieved through the speaker's rehearsed delivery during recording: no post-hoc audio manipulation was applied. The natural prosody, intonation and pause structure of the recording were therefore preserved. At 86 WPM, delivery falls below typical conference speech rates, which generally range from 100 to 120 WPM (Riccardi 2015). The reduced rate was a deliberate methodological choice. The study aimed to isolate the effect of ASR support on numerical rendition. Faster conference-typical rates (100–120 WPM) could have introduced a confounding variable, such as rate-induced cognitive overload. Holding speech rate constant at 86 WPM ensures that observed between-condition differences reflect the contribution of ASR support specifically. The ecological validity trade-off is acknowledged as a limitation and revisited in the Discussion.

Both speeches contained 59 numerical items. As with speech rate, this was a product of selection criteria rather than coincidence—speeches were chosen from a larger pool on the basis of comparable numerical density, topic neutrality relative to participant background knowledge and suitability for the target language pair.

Speech 1 (Control Condition): Topic: Global Warming. Duration: 9 minutes and 16 seconds. All the participants interpreted this speech without any technological support.

Speech 2 (Experimental Condition): Topic: E-commerce. Duration: 7 minutes and 47 seconds. All the participants interpreted this speech with simulated ASR support—a pre-recorded video display in which numerical information appeared on screen synchronized with the speaker’s delivery. The display was not generated by a live ASR system; the numerical information was prepared in advance and assembled in iMovie as a pre-recorded video in which each figure appeared as white text on a black background, timed to appear on screen as the speaker articulated it. Prior to the experimental sessions, participants were provided with a terminology bank covering the lexical and topical content of both speeches, in line with standard interpreter training practice for advance preparation.

Numerical categorization

Before analysis could begin, all numerical items in both source texts were identified and classified. This step was necessary because not all numbers present the same cognitive challenge—a single-digit figure and a seven-digit figure make fundamentally different demands on the interpreter and collapsing them into a single category would obscure meaningful variation in the results.

The numbers appearing in the two source texts were grouped into five categories reflecting the types actually present in the speeches. This typology is informed by Jones's (2002) account of the processing dimensions that make numbers cognitively demanding—arithmetic value, order of magnitude, unit, extra-linguistic referent and relative value—on the premise that different number types draw on these dimensions to different degrees and therefore present distinct processing demands. Rather than adopting an existing scheme in full, the categories were selected to capture the specific items that occurred in this study. The resulting five-category scheme is as follows:

High-magnitude numbers refer to figures with four or more digits: thousands, millions, billions. These are consistently identified in prior literature among the most cognitively demanding numerical types (Mazza 2001; Pinochi 2009; Desmet et al. 2018).

Low-magnitude numbers cover single and double-digit whole numbers. These are generally less demanding but can still produce phonemic errors in English, particularly between near-homophones such as “thirteen” and

“thirty.” Dates include specific calendar years and historical references. Although structurally simple, dates are unforgiving—an error of even one digit changes the meaning entirely and there is no contextual redundancy to help the listener reconstruct the correct figure.

Percentages are numerical values expressed as a proportion of 100. They appear frequently in both speeches and carry high informational density relative to their length.

Decimals are non-percentage figures involving a decimal point. These proved to be the most resistant to improvement even with technological support, for reasons discussed in the results section.

Each numerical item in both source texts was assigned to one of these categories prior to analysis, producing a structured inventory against which the interpreted output could be systematically compared.

Error typology and data analysis

The interpreted recordings were transcribed in full and coded against the source text on a number-by-number basis. Each numerical item was first evaluated as either correct or incorrect. Items judged incorrect were then assigned to one of four error categories adapted from the typology developed by Pinochi (2009), which itself builds on the earlier framework of Braun and Clarici (1996). Following Pinochi, a lexical error preserves the order of magnitude while corrupting one or more digits (e.g., 277,000 → 276,000), whereas a syntactical error renders the wrong order of magnitude even when the figures themselves are correct (e.g., 47,000 → 47 million). This distinction is consistent with Frittella’s (2019) coding of numerical errors. For the English–Turkish pair specifically, Bozok and Kınca (2022) document numbers as a recurrent error source among student interpreters, situating omission within Pym’s risk framework. Pinochi’s original typology contains seven categories; three were excluded from the present study. The miscellaneous category was omitted on the grounds that unclassifiable errors cannot be meaningfully interpreted within either Gile’s or Seeber’s theoretical frameworks. Approximation and transposition, while well-documented in the literature, were not observed in either condition of this study and were therefore excluded from the analysis. The remaining four categories are retained as defined:

Omission occurs when a numerical item is entirely absent from the target output, or replaced by a non-specific expression such as “many” or “some.” This is the most common error type in unsupported interpreting and functions, under Gile’s (1995) framework, as a tactical sacrifice—the interpreter drops the number to preserve the coherence of the surrounding speech. Although it is generally considered a tactical sacrifice, according to Pym (2008), interpreters also make omissions based on their evaluation of the communicative risk. Thus, low risk omissions “are part of a general economy

of time management, mostly as part of a general strategy of implicitation” (Pym 2008: 95). Therefore, “interpreters might make omissions as a time-saving strategy for pragmatic reasons unless the omitted information is crucial for the purposes of the communicative act”, which requires considering the contextual dimension of omissions besides the cognitive (Kincal 2020: 99).

The concept of risk management in translation (including interpreting) by Anthony Pym is highly relevant in this context (Pym 2025). With a broader view of interpreting as a form of intercultural communication, the strategies that orient the decisions of interpreters from cognitive to social are defined by their intuitive assessments of the “effects that each decision will have on the imagined receiver”, which can be considered as “embodiments of successful risk management decisions” (2025: 3). Thus, numbers can be sacrificed as a risk management strategy rather than a high-risk error.

Lexical error occurs when the order of magnitude is correct but specific digits or components are changed—“346” rendered as “436”, for example. These errors suggest that the interpreter captured the scale of the number but lost precision at the digit level.

Syntactical error occurs when the digits are correct, but the order of magnitude is rendered incorrectly—“47” produced as “470”, or “47 million” instead of “47 thousand.” In the English–Turkish context, this error type is particularly significant because the two languages differ in how they structure large numbers, creating a production-level challenge that persists even when the source figure has been correctly perceived.

Error of phonemic perception results from the phonological similarity of certain English figures—most commonly the -teen/-ty distinction, as in “fourteen” versus “forty.” These errors are a source feature of the English input rather than a target language production failure.

Results

This section presents the findings of the comparative analysis between the unsupported condition (Speech 1) and the ASR-supported condition (Speech 2). Results are organized across three dimensions: global accuracy and individual performance, accuracy gains by numerical category and the shift in error profiles between conditions.

Accuracy and individual performance

Across all six participants, accuracy improved in the ASR-supported condition without exception. In Speech 1, individual accuracy rates ranged from 54.2% (INT5) to 76.3% (INT2), with a cohort mean of 64.4%. In Speech 2, the range shifted upward from 80.0% (INT3) to 96.4% (INT1), with a cohort mean of 87.6%. The magnitude of individual improvement varied

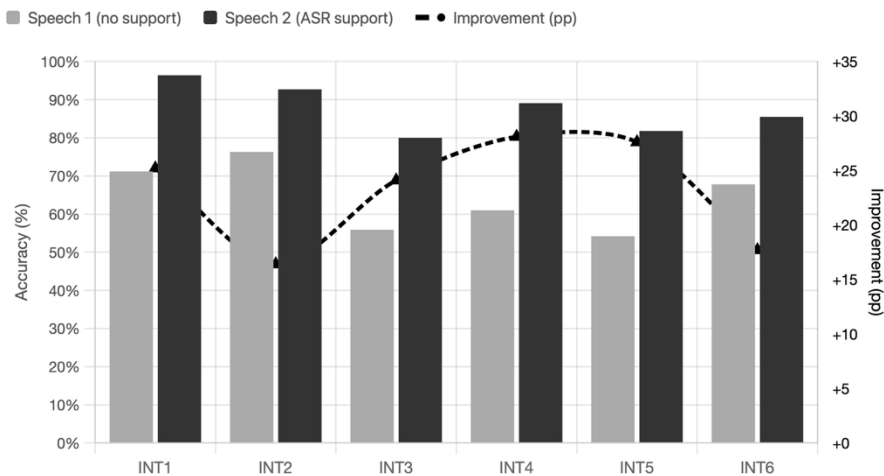
considerably. INT2, who entered the experimental condition with the highest baseline, showed the smallest gain at 16.5 percentage points. INT4, by contrast, improved by 28.1 points—the largest individual gain in the cohort. INT1, INT3 and INT5 showed improvements of 25.2, 24.1 and 27.6 points respectively, clustering around a similar range. INT6 improved by 17.7 points, placing alongside INT2 at the lower end of the improvement distribution.

The individual figures are as follows:

Table 1: Accuracy and individual performance

	Speech Accuracy 1	Speech Accuracy 2	Improvement
INT1	71.2%	96.4%	+25.2 pts
INT2	76.3%	92.7%	+16.4 pts
INT3	55.9%	80%	+24.1 pts
INT4	61%	89.1%	+28.1 pts
INT5	54.2%	81.8%	+27.6 pts
INT6	67.8%	85.5%	+17.7 pts

Figure 1: Individual accuracy rates (%) in Speech 1 and Speech 2, with improvement in percentage points (right axis).



Two observations are worth noting at this stage. First, the two participants with the smallest gains—INT2 and INT6—were also among those with the higher baselines, suggesting a ceiling effect may have constrained their improvement. Second, the participants with the lowest baselines showed the steepest gains, indicating that the visual support was especially consequential for those who struggled most in the unsupported condition.

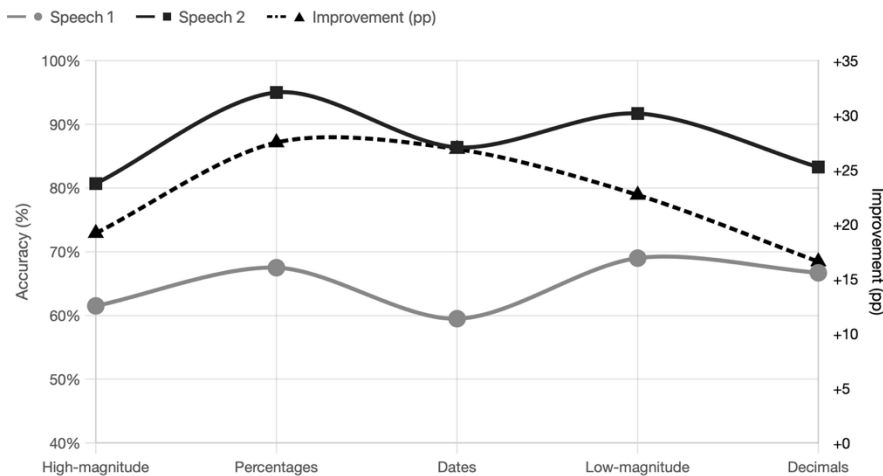
Accuracy gains by numerical category

The improvement was not uniform across numerical types. Table 2 presents the cohort-level accuracy figures for each category across both conditions.

Table 2: Accuracy gains by numerical category

Number Type	Speech Accuracy 1	Speech Accuracy 2	Improvement
High Magnitude	61.5%	80.7%	+19.2 pts
Percentages	67.5%	95%	+27.5 pts
Dates	59.5%	86.4%	+26.9 pts
Low Magnitude	69%	91.7%	+22.7 pts
Decimals	66.7%	83.3%	+16.6 pts

Figure 2: Cohort-level accuracy (%) by numerical category across both conditions, with improvement in percentage points (right axis).



Percentages showed the largest absolute gain at 27.5 percentage points, reaching a near-ceiling accuracy of 95.0% in the supported condition. Dates followed closely with a 26.8-point improvement. Both categories share a structural feature that likely explains their responsiveness to visual support: they are high in informational density but relatively standard in format, meaning that once the figure is visible, its integration into the target output is comparatively straightforward.

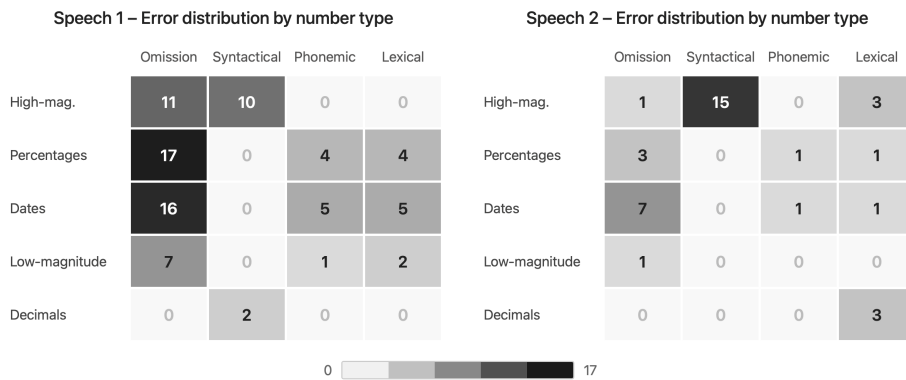
High-magnitude numbers showed a more modest improvement of 19.2 points, remaining the lowest-accuracy category in both conditions. This is significant: despite the visual display making the figure available, participants still struggled with these items more than any other types. Decimals showed

the smallest gain at 16.7 points, suggesting that even with visual support, this category presents particular challenges—a point returned to in the discussion.

Error profile shift

The heatmap data reveal not just a reduction in errors but a fundamental change in the type of errors produced, which is arguably the more theoretically significant finding. Notably, approximation—a strategy well-documented in the literature—was effectively absent in both conditions, suggesting that the participants in this study either committed to a full rendition or dropped the number entirely.

Figure 3: Error distribution by number type and error category for Speech 1 (left) and Speech 2 (right).



Note: Cell values indicate total error count across all six participants. Darker shading reflects higher frequency.

In Speech 1, omission was the dominant error type across virtually all categories. Percentages produced 17 omissions, Dates 16 and High-magnitude numbers 11—together accounting for the large majority of all errors in the unsupported condition. Syntactical errors were present but secondary, concentrated in High-magnitude numbers (10 instances). Phonemic errors appeared in moderate numbers across Percentages, Dates and High-magnitude items. Lexical errors were confined to Percentages, Dates and Low-magnitude numbers, with 4, 5 and 2 instances respectively.

In Speech 2, this profile changed substantially. Omissions decreased substantially across most categories: Percentages fell from 17 to 3, Dates from 16 to 7 and High-magnitude numbers from 11 to 1. Phonemic errors similarly declined. However, syntactical errors in the High-magnitude category increased sharply, rising from 10 to 15—the only error type to worsen in the supported condition. Lexical errors in the High-magnitude category also rose slightly, from 0 to 3.

The pattern is clear: ASR support largely solved the perceptual problem. Participants who previously dropped numbers entirely were now capturing them from the visual display. What the support did not resolve—and in some respects made more visible—was the production problem. With the figure now in hand, participants still had to construct it in Turkish, and it is at this structural integration stage that errors concentrated in the supported condition.

Discussion

The results of this study invite a reconsideration of how cognitive load is theorized in simultaneous interpreting—not a wholesale rejection of existing frameworks, but a refinement of what they can and cannot explain when technological support is introduced into the interpreting task at hand.

Gile's Effort Model remains a productive starting point. The baseline data from Speech 1 maps cleanly onto its central prediction: when the combined demands of Listening, Memory and Production exceed available capacity, the system sheds load by the path of least resistance. In numerically dense speech, that path is omission. The high omission rates observed across all participants in unsupported condition, 17 instances for percentages alone, 16 for dates, are precisely what Gile's framework would lead us to expect. Interpreters were not making careless errors; they were making triage decisions under pressure, sacrificing individual figures to preserve the coherence of the surrounding speech—what Pym (2025) characterizes as the management of communicative risk.

Where the model shows its limits is in accounting for what happened next. Under a strictly unitary resource view, introducing a visual display during interpreting should add to the cognitive burden, another input stream competing for the same finite pool of attention. The data suggests otherwise. Accuracy improved universally across all six participants and omissions collapsed in every category where the visual support was available. These results suggest the visual support relieved cognitive load rather than adding to it. This is where Seeber's (2011) Cognitive Load Model offers a more adequate account. If visual-spatial and auditory-verbal processing draw on partially independent resource pools, as Wickens' (2002) Multiple Resource Theory proposes, then a congruent visual display does not compete with the auditory stream—it complements it. The number on screen reaches the interpreter through a different channel than the number in the audio and the two together are less costly than either alone under saturating conditions.

That said, the data also complicates any straightforward celebration of multimodal support, and this is where the study's most significant finding lies.

The shift in error profile between Speech 1 and Speech 2 is not merely a quantitative improvement—it is a qualitative transformation in the nature of failure. In unsupported condition, the dominant error was omission: the

interpreter did not have the number and could not produce it. In the supported condition, omission was largely eliminated, but syntactical errors in the high-magnitude category increased from 10 to 15 instances. Interpreters now had the number—they could see it—but they still struggled to render it correctly in Turkish. The problem had moved downstream, from perception to production.

This distinction matters theoretically. Gile's model treats cognitive saturation as a single phenomenon; the data suggests it has at least two distinct stages. The first is perceptual saturation, the failure to capture a stimulus under auditory overload. Technology, in this study, effectively solved that problem. The second is structural saturation, the failure to integrate a correctly perceived stimulus into the morphosyntactic architecture of the target language under time pressure. Technology did not solve that problem and arguably made it more visible by removing the perceptual problem that had previously masked it.

The structural challenge is particularly acute in English–Turkish interpreting, where production failures can arise independently of perception. The clearest case in the present data involves the order of magnitude: English “million” and “billion” correspond to Turkish “*milyon*” and “*milyar*”, near-identical forms that compete under time pressure. When an interpreter sees “2.3 billion” on screen, the perceptual task is complete—the figure has been correctly received—yet selecting the right scale word in production remains demanding, and an error here relocates the entire figure across magnitudes while leaving the digits intact. The persistence of syntactical errors in the supported condition suggests that this production-level bottleneck is not addressable through visual input alone: seeing the number does not resolve the competition between scale words that arises at the moment of production.

The category-level results add further nuance to the multimodal facilitation picture. Percentages and dates showed the largest gains—27.5 and 26.8 percentage points respectively—while decimals showed the smallest at 16.7 points. This variation is not random. Percentages and dates are structurally simple in Turkish: once the figure is captured, its integration into the target output is relatively formulaic. Decimals, by contrast, involve notational complexity that requires cognitive investment at the production stage regardless of how the figure was perceived. The visual display helps most where the production task is straightforward; it helps least where production itself is the bottleneck. This is consistent with Seeber's (2017) observation that visual input functions as a cognitive aid specifically when it offloads a process that would otherwise compete with production—not as a universal solution to all sources of difficulty.

Taken together, these findings suggest a more granular picture of the “augmented interpreter” than the literature has so far offered (Fantinuoli and Dastyar 2022). Technology does not simply reduce cognitive load—it redistributes it. The interpreters who work with ASR support are no longer

primarily managing a perceptual problem; they are managing a structural one. This is a meaningful shift, but it is not a resolution. The booth becomes a different cognitive environment, not an easier one.

Conclusion

The findings of this study converge on a point that is both empirically grounded and theoretically generative: technology shifts the interpreter's failure mode rather than eliminating it. This is not a pessimistic conclusion, the accuracy gains were real and consistent, but it is a clarifying one.

The baseline condition confirmed what the literature on numerical processing in SI has long established. Numbers are among the most frequent sources of failure in interpreting performed without in-booth technological assistance and omission is the dominant response - in some cases, the strategy - to that failure. Student interpreters working without technological assistance sacrificed figures at high rates, particularly percentages, dates and high-magnitude numbers, not out of incompetence but out of rational prioritization under pressure. Gile's Effort Model describes this dynamic well and the Speech 1 data offer straightforward empirical support for its core prediction.

The introduction of simulated ASR support changed the picture substantially. Accuracy improved for every participant without exception and omissions declined sharply across nearly all numerical categories. The visual display allowed interpreters to bypass the auditory bottleneck that had previously made numerical capture so costly. This is the outcome Seeber's multimodal framework predicts and the consistency of the effect across participants and categories gives it credibility even within the constraints of a small exploratory sample.

But the more significant finding is what happened to the nature of failure. With the perceptual problem largely resolved by visual support, structural errors became the dominant failure mode. Interpreters who could now see the figure still struggled to render high-magnitude numbers correctly in Turkish, where the syntactic demands of production remain considerable regardless of how the stimulus was received. Technology solved one bottleneck and exposed another.

This points toward something the field has perhaps underappreciated in its enthusiasm for AI-assisted interpreting. Aggregated accuracy metrics, the kind routinely reported in CAI studies, can tell us what the results of technological intervention are, but they are less equipped to tell us how those results come about, or at what cognitive cost. The distinction between Gile and Seeber is not merely a theoretical disagreement; it is a methodological one. A unitary resource model and a multimodal model make different predictions about where failure will concentrate, and those predictions have practical consequences for how we design tools, evaluate performance and

train interpreters. Without that theoretical granularity, we risk building an “augmented interpreter” model on foundations that do not fully account for the cognitive architecture of the task.

The English–Turkish language pair has received comparatively little attention in prior literature (apart from Arzik Erzurumlu and Demir 2022) and this study is exploratory in scope. But it suggests that the pair presents specific challenges, particularly in the structural integration of high-magnitude numbers, that deserve dedicated investigation. More broadly, it adds to the growing body of evidence that simultaneous interpreting cannot be adequately understood through auditory-verbal models alone. The booth has always been a multimodal environment; what AI-backed technologies have done is make that fact impossible to ignore.

AI acts as a facilitator; yet the degree to which it does so and the cognitive costs it redistributes rather than removes, remain open questions that warrant further investigation.

Declaration of AI use

In preparing this manuscript, the authors used Claude for preparing the tables. All AI-assisted content was reviewed, edited and verified by the authors, who take full responsibility for the accuracy and integrity of the final text.

References

- Arzik Erzurumlu, Özüm, and Demir, Perihan. 2022. “Taking stock of the COVID-19 working conditions on the performance of the interpreters: Rendering the numbers in the 2020 American Presidential Debates”. *Journal of Translation Studies* 32: 25–54.
- Baigorri-Jalón, Jesús. 2014. *From Paris to Nuremberg: The Birth of Conference Interpreting*. Translated and edited by Holly Mikkelson and Barry Slaughter Olsen. Amsterdam and Philadelphia: John Benjamins.
- Bozok, Nazlıgül, and Kıncal, Şeyda. 2022. “Strategies and errors in simultaneous interpreting: A student-oriented experiment in English–Turkish language pair”. *Current Trends in Translation Teaching and Learning E*, 32–75. <https://doi.org/10.51287/cttl20222>
- Braun, Sabine and Clarici, Andrea. 1996. “Inaccuracy for numerals in simultaneous interpretation: Neurolinguistic and neuropsychological perspectives”. *The Interpreters' Newsletter* 7: 85–102.
- Defrancq, Bart, and Fantinuoli, Claudio. 2021. “Automatic speech recognition in the booth: Assessment of system performance, interpreters’ performances and interactions in the context of numbers”. *Target* 33(1): 73–102.

- Dehaene, Stanislas, Spelke, Elizabeth S., Pinel, Philippe, Stanescu, Ruxandra, and Tsivkin, Sanna. 1999. "Sources of mathematical thinking: Behavioral and brain-imaging evidence". *Science*, 284(5416): 970–974. <https://doi.org/10.1126/science.284.5416.970>
- Desmet, Bart, Vandierendonck, Mieke, and Defrancq, Bart. 2018. "Simultaneous interpretation of numbers and the impact of technological support". In C. Fantinuoli (Ed.), *Interpreting and Technology*. Berlin: Language Science Press. 13-27.
- Fantinuoli, Claudio. 2018. "Interpreting and technology: The upcoming technological turn". In C. Fantinuoli (Ed.), *Interpreting and Technology*. Berlin: Language Science Press. 1-12.
- Fantinuoli, Claudio and Dastyar, Vorya. 2022. "Interpreting and the emerging augmented paradigm". *Interpreting and Society: An Interdisciplinary Journal* 2(2): 185–194.
- Frittella, Francesca Maria. 2019. "70.6 billion world citizens': Investigating the difficulty of interpreting numbers". *Translation & Interpreting* 11 (1), 79–99. <https://doi.org/10.12807/ti.111201.2019.a05>
- Gile, Daniel. 1995. *Basic Concepts and Models for Interpreter and Translator Training*. Amsterdam and Philadelphia: John Benjamins.
- Gile, Daniel. 2009. *Basic Concepts and Models for Interpreter and Translator Training* (Rev. ed.). Amsterdam and Philadelphia: John Benjamins.
- Jones, Roderick. 2002. *Conference Interpreting Explained*. Manchester: St. Jerome Publishing.
- Kıncal, Şeyda. 2020. "The impact of non-native English on omissions in simultaneous interpreting". *Journal of Translation Studies* 28: 98-109.
- Korpala, Paweł, and Stachowiak-Szymczak, Katarzyna. 2018. "The whole picture: Processing of numbers and their context in simultaneous interpreting". *Poznan Studies in Contemporary Linguistics* 54(3): 335–354.
- Lu, Xinchao, and Fantinuoli, Claudio. 2025. "Machine and computer-assisted interpreting: Innovations in and implications for interpreting practice, pedagogy and research". *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 24. <https://doi.org/10.52034/lans-tts.v24i.869>
- Mazza, Cristina. 2001. "Numbers in simultaneous interpretation". *The Interpreters' Newsletter* 11: 87–104.
- Mead, Peter. 2015. "Numbers". In Franz Pöchhacker (ed.), *Routledge Encyclopedia of Interpreting Studies*, 286–288. Abingdon and New York: Routledge.
- Pinochi, D. 2009. Simultaneous Interpretation of Numbers: Comparing German and English to Italian. An Experimental Study. *The Interpreters' Newsletter*, 14, 33–57. <http://hdl.handle.net/10077/3465>
- Pöchhacker, Franz. 2015. *Introducing Interpreting Studies* (2nd ed.). New York and London: Routledge.

- Pym, Anthony. 2008. "On omission in simultaneous interpreting: Risk analysis of a hidden effort". In G. Hansen et al. (eds) *Efforts and Models in Interpreting and Translation Research: A Tribute to Daniel Gile*. Amsterdam and Philadelphia: John Benjamins. 83-105.
- Pym, Anthony. 2025. *Risk Management in Translation*. Cambridge: Cambridge University Press.
- Riccardi, Alessandra. 2015. "Speech rate". In F. Pöchhacker (Ed.), *Routledge Encyclopedia of Interpreting Studies*. London: Routledge. 397–399.
- Seeber, Kilian G. 2011. "Cognitive load in simultaneous interpreting: Existing theories—New models". *Interpreting* 13(2): 176–204.
- Seeber, Kilian G. 2012. "Multimodal input in simultaneous interpreting: An eye-tracking experiment". In L. N. Zybatow, A. Petrova, and M. Ustaszewski (Eds.), *Proceedings of the 1st International Conference TRANSLATA*. Frankfurt am Main: Peter Lang. 341–347.
- Seeber, Kilian G. 2017. "Multimodal processing in simultaneous interpreting: A multiresource model". In J. W. Schwieter, A. Ferreira (eds), *The Handbook of Translation and Cognition*. Hoboken, New York: Wiley. 1-16.
- Seeber, Kilian G., Keller, Laura, and Hervais-Adelman, Alexis. 2020. "When the ear leads the eye – the use of text during simultaneous interpretation". *Language, Cognition and Neuroscience* 35(10): 1480–1494.
- Timarová, Šárka. 2012. *Working Memory in Conference Simultaneous Interpreting*. Unpublished doctoral dissertation. Charles University/KU Leuven.
- United Nations General Assembly. 1947. Resolution 152 (II): Simultaneous Interpretation. A/RES/152(II), 15 November 1947.
- Wickens, Christopher D. 2002. "Multiple resources and performance prediction". *Theoretical Issues in Ergonomics Science* 3(2), 159–177.
- Yuan, Lu, and Wang, Binhua. 2023. "Cognitive processing of the extra visual layer of live captioning in simultaneous interpreting: Triangulation of eye-tracked process and performance data". *Ampersand* 11, 100131.

Quality in multilingual crisis communication: a reception-oriented perspective

YU HAO

University of Melbourne, Australia

Multilingual crisis communication relies on timely and reliable translation, yet what constitutes a “good” translation remains conceptually underexplored in this high-stakes, culturally loaded field. This position paper reviews previous studies and argues “quality” should be treated not only as a text’s intrinsic property but as real-world communicative success. More specifically, this position paper conceptualizes quality as a series of effect-related parameters: the degree to which intended audiences can access messages in time (“timeliness”), make sense of them (“understandability”), trust them as credible and relevant (“trustworthiness”) and act on them appropriately (“actionability”).

Keywords: crisis communication; behavioral-change emergency messaging; quality assessment; reception-oriented perspective

Introduction

Crisis translation is often regarded as a matter of language justice and equity (e.g., O’Brien 2022). Unequal and delayed access to reliable translated information can exacerbate existing vulnerabilities and create new risks for public safety during health crises and other emergencies including earthquakes, bushfires and floods. Thus, a major challenge faced by multicultural societies in crisis and emergency situations is the need to provide reliable translated messages so that all communities have equitable access to critical information in their own languages. This challenge is particularly acute given that crisis communication constitutes a high-stakes genre whose purpose is not only to inform but also to elicit appropriate behavioral responses in a timely manner; not to mention the linguistic and cultural nuances involved when engaging with divergent communities that experience barriers in accessing official messaging, which is often initially prepared in the society’s dominant language.

Over the years, studies have examined translation in these high-stakes communicative scenarios (e.g., studies reviewed in Federici 2020; Todorova, 2024). To mitigate the challenge of multicultural crisis translation, some studies focus on the training of non-professional citizen translators (e.g., Federici and Cadwell 2018), which is important in situations where professional translation is either unavailable (in the cases of low-resource languages) or overwhelmed due to limited human resources. Bilingual/bicultural members of a community can thus be trained as citizen translators to create locally available and well-adapted services. This challenge is also being met with advances in machine translation and generative AI systems that automate or augment communication across various domains. In this context, recent studies in this field have explored the use of fully automated translation (e.g., Hajek et al. 2024; do Carmo 2025) and semi-autonomous workflows involving both human and non-human agents (e.g., Hao and Pym 2026). Others examine “trust” in multilingual communication, which they view as a layered and socially mediated construct. On this view, linguistic clarity seems to enable initial comprehension while culturally grounded relationships and credible mediators sustain the desirable behavioral changes (e.g., Macreadie et al. 2025).

A central question merits careful analysis: how do we know whether a translated crisis-related message is good enough? Although “quality” is addressed either directly or indirectly in many studies on crisis communication, there remains limited conceptual clarity with respect to what constitutes “good” translation in public emergencies. This becomes increasingly relevant, as translation continues to shift from static text production to dynamic and increasingly iterative human-AI interactions. This position paper, therefore, reviews and examines how “translation quality” has been constructed in crisis messaging, with special attention to the shift from a purely textual evaluation towards a reception-oriented framework.

Translation quality redefined: Why reader response matters?

This section presents a conceptual discussion of the theories over the years, elaborating on what grounds we should judge the “success” of translation and why textual comparison alone could be reductive.

Although the equivalence paradigm generally privileges source-text resemblance, Nida (1964) contends that translation should also aim to produce “the same effect on the reader as the original text did” (Nida 1964: 159). While this dynamic or functional equivalence perspective has been criticized as reductive by those with a *Skopos*-oriented mindset (e.g., Vermeer 1989), since translations may aim to achieve effects different from those of the source text, it nonetheless highlights the need to factor in how target readers might understand and respond to a translation.

As shown in the *Skopos* paradigm (Vermeer 1989), the evaluative center of gravity has shifted away from retrospective textual comparison between translation and its start text, as anchored in the equivalence paradigm (see more detailed discussion in Pym 2023). As *Skopos* theory allows that a single text can be legitimately translated in different ways, each yielding a valid translation (Reiss and Vermeer 1984; Vermeer 1989), “quality” is inherently purpose- and readership-dependent. Thus, quality assessment models that prioritize formal proximity to the original could risk mistaking surface-level mimicry for high quality, even when the intended communicative purpose might not be achieved. Also according to Vermeer’s *Skopos* rules (Reiss and Vermeer 1984), the purposes that could govern all subsequent translation decisions might be defined by various actors who expect value from translation, including the commissioner’s rationale (Nord 2005), the reader’s intentions (Vermeer 1989) and/or what the trained translator judges the purpose to be (Vermeer 1989; Holz-Mänttari 1984). In these views, readers and commissioners are made epistemically important: if they have a say in what translation should be made *for*, then empirical evidence should be gathered on whether translation functions effectively for the end users in the intended context.

Further, norm-based theories, as proposed by Toury (1995/2012) based on sociological concepts developed by Itamar Even-Zohar and earlier sociologists, help explain why readers’ sense of “normal” matters in quality assessment. In descriptive translation studies (DTS), translation is characterized as norm-governed behavior shaped by the conventions and expectations of the target culture (Toury 1995/2012). That is, a translation’s uptake and legitimacy are mediated by shared norms constructed by communities of readers, critics, publishers and other gatekeepers at particular historical moments and within specific cultural contexts. In this sense, any account of “quality” must remain grounded in the cultural and historical conditions under which translations are produced, circulated and interpreted. Moreover, Toury’s polarized distinction between adequate and acceptable translation (Toury 1980) is pertinent here: whereas an adequate translation adheres to source-text norms, an acceptable translation aligns with target-side norms to achieve naturalness and credibility. So, what counts as “good” depends only partly on its relation to the source text and perhaps more predominantly on how it positions itself *vis-à-vis* the recipient culture’s expectations of well-formed and credible texts.

Chesterman (1998) further expands the effects-orientation in Nida’s (1964) position on dynamic equivalence into a conceptual framework which differentiates “proximate (cognitive and affective) effects”, “behavioral effects” and “tertiary effects.” He argues that translation quality should be assessed through empirical investigation of its effects, including its impact on readers’ mental states, subsequent behaviors and macro-level cultural change, such as shifts in norms, genres and ideological repertoires. All of those cannot

be captured through sentence-level comparison alone. Chesterman's three "laws" further emphasize the dynamic status of translation effects: translations may produce different effects on different readers ("law of heterogeneous effect"); effects may vary over time even for the same reader ("law of changing effect"); and translations often generate multiple effects simultaneously ("law of multiple effects") (Chesterman, 1998). If such effects are heterogeneous and temporally unstable, this raises an important question for translation quality assessment: for whom, under what conditions can a translation be considered "good"? Beyond textual analysis, it therefore requires us to examine empirically how people actually understand and appreciate the translated text, as well as the communicative conditions under which translations are received, interpreted and acted upon.

Recent studies on quality in multilingual crisis communication

This section then reviews how translation quality has been conceptualized and assessed in existing literature on multilingual communication, particularly in healthcare and emergency settings. Previous literature shows divergent ontological positions with respect to what constitutes "quality," which in turn give rise to two broad approaches to its empirical investigation: (1) an output-focused approach, grounded in the textual and linguistic analysis of translations and (2) a reception-oriented approach, which prioritizes how recipients understand, interpret, respond to and act upon the translation. Within the first approach, quality is treated as an intrinsic property of the translated text and is often evaluated through manual error analysis of the translated text or automatic evaluation metrics (e.g., BLEU, METEOR). While the reception-oriented approach conceptualizes quality in terms of communicative outcomes and assesses quality through human-centered protocols in addition to textual analysis. This study hastens to add that these two approaches are not mutually exclusive but intertwined. Textual analysis could nevertheless be conducted by the researchers in preparation for eliciting any empirical evidence from the authentic recipients (e.g., Hajek et al. 2024; Hao and Pym 2026). Some reception studies may also invite bilingual (non-)professionals to compare the translation against its original (e.g., O'Brien et al. 2018) or in relation to the target language conventions.

Output-focused quality: textual analysis and error taxonomies

Within the output-focused paradigm, translation quality is conceptualized as a measurable attribute of the translated artefact itself. This approach has been applied to the evaluation of outputs produced from scratch, generated fully autonomously, or developed through hybrid workflows, through the analysis of linguistic features with reference to start-text fidelity and target-language

norms of fluency and naturalness. These analyses of quality seem predominantly negative and reductive, grounded in an error-based, deficit-oriented mindset (e.g., Staiano et al. 2025; Guo 2026). That is, quality is inferred from deviations from an assumed ideal: the presence, severity and distribution of errors (as in Multidimensional Quality Metrics, MQM framework) serve as indicators of lower quality: quality scores decrease as the number and gravity of errors increase. Textual analysis could either be deductively based on existing error taxonomies or proceed inductively through observing recurring error patterns in the texts at hand.

Over the years, this output-focused approach has been adopted to examine the feasibility of machine-assisted crisis translation. For instance, O'Brien et al. (2019) compared three translation modes: from-scratch translation, raw machine translation and post-edited translation. In this study, professional evaluators and bilingual participants were asked to assess these outputs in terms of adequacy (adherence to source meaning) and fluency (linguistic well-formedness), often supplemented by error annotation procedures. The aim was to identify systematic errors in machine translation in crisis contexts, particularly for low-resource language pairs.

Similarly, Pym et al. (2022) conducted a detailed textual analysis of raw machine translation in public health communication settings. Their study highlights the risks of using raw machine output in high-stakes scenarios, which identified recurrent error patterns such as untranslated text in images, omissions and ambiguity in context-dependent terms and syntactic structures. On top of these studies, this output-focused mindset also seems central to the training of bilingual mediators in crisis settings. For example, Federici et al. (2019) highlight the need for both professional and citizen mediators to maintain terminological consistency, domain-specific accuracy and stylistic appropriateness in their output, particularly when dealing with risk-sensitive content.

However, the output-focused perspective can potentially be problematic, as it tends to treat semantic accuracy and naturalness alone as the primary indicators of quality. In this sense, this reductive view seems to equate quality with communicative success in crisis messaging, assuming that textual fidelity and naturalness will always lead to clear understanding and desirable behavioral changes among its recipients. Researchers in this area have increasingly recognized the limitations of relying solely on this approach for quality assessment. As O'Brien and Federici (2020) argue, crisis communication often involves multimodal, time-sensitive and culturally embedded messages, in which linguistic correctness alone may not be sufficient to guarantee effective communication.

Reception-oriented quality: Understandability, actionability, trustworthiness and timeliness

A growing body of research has advocated for a reception-oriented perspective to translation quality (e.g., O'Brien et al. 2018; Hajek et al. 2022). Within this paradigm, quality is no longer merely treated as an inherent attribute of the translated text, but also as an emergent outcome from the interactions between the text, its prospective readers and the broader communicative context.

Recent studies reflect a broader epistemological shift towards evaluating quality as a context-dependent phenomenon, grounded in user experience and communicative outcomes. Beyond textual analysis, these reception studies all empirically examine how culturally, socially and age-heterogeneous groups understand and respond to emergency messages in various settings, although each has paid particular attention to certain aspects. For instance, O'Brien and Cadwell (2017) conducted a series of comprehension tests in a case study in Kenya, to assess the extent to which original and translated messages are understood by prospective readers. The translated public health messaging in Kiswahili was found more likely to elicit desirable sense-making than the English original.

In parallel, many have argued that in crisis contexts, where timely and accurate action can be lifesaving, the translation's ability to elicit desirable behavioral changes should also be prioritized. For instance, Ogie and Perez (2020) compared community feedback on intended behavioral responses to texts translated by professional translators and citizen mediators. Their findings indicate that professionally translated texts may fail when the source text itself is ambiguous or overly technical. Responses to translated emergency messages depend largely on clarity, simplicity and cultural resonance, areas in which citizen translation seems to outperform its counterpart.

Recent studies have also factored in time-based evaluation, measured in seconds per word and task completion time. These measures are particularly pertinent in crisis and public health contexts, where timely messaging could profoundly influence communicative impact. Working with automated translation (such as post-editing machine or AI output) has traditionally been associated with speed gains compared with from-scratch translation (e.g., O'Brien 2011; Plitt and Masselot 2010). However, previous studies (e.g., Pym et al. 2022) also showed that raw automated output should not be used in crisis translation despite the faster turnaround, given that even if the output is 99% accurate, the remaining 1% of errors could potentially lead to fatal consequences during emergencies. In other words, speed should not be traded off against reliable, trustworthy messages and overall communicative effectiveness.

Lastly, this shift in focus also emphasizes trustworthiness as an evaluative dimension, particularly in high-stakes contexts where credibility can potentially shape how recipients respond to the message. For instance, Pym et al. (2022) focused on this dimension in a COVID-19 public messaging context, using qualitative and observational methods. Trust is inferred indirectly from cases of distrust, such as translation errors, delayed or inaccessible messaging and from communities' reliance on alternative mediators or information sources other than the official government translation. Macreadie et al. (2025) further argue that translation is not only about conveying an accurate message but about building trust between institutions and audiences in the context of cross-cultural vaccination messaging. The study distinguishes between "thin trust" (based on surface accuracy and institutional authority) and "thick trust" (grounded in cultural relevance and social relationships) and reports that most institutional translations achieve only the former. That is, even linguistically accurate translations can sometimes fail in public health messaging, especially if audiences do not perceive the text as culturally trustworthy. More recent studies in this field on human-centered AI also introduce evaluation criteria such as trustworthiness, safety and user control (e.g., Briva-Iglesias and O'Brien 2026).

Drawing on the literature, this position paper argues that, beyond linguistic correctness, translation quality in crisis settings should be evaluated in terms of its "effects" along four dimensions. Each dimension should be empirically tested with authentic message recipients in context.

- Understandability: the degree to which receivers can make sense of the text (e.g., in O'Brien and Cadwell 2017; Rossetti 2019; Alexander and Pescaroli 2020; Ogie and Perez 2020);
- Actionability: the extent to which receivers are attentive to information and make desirable behavioral changes (e.g., Sengupta et al. 2024)
- Trustworthiness: whether receivers perceive the message as credible (e.g., O'Brien 2022; Pym et al. 2022; Seale et al. 2024; Macreadie et al. 2025) and lastly
- Timeliness: whether receivers can access up-to-date information through convenient channels in a timely manner (O'Mara and Carey 2019; Hajek et al. 2024).

These four dimensions are not mutually exclusive and can interact with one another in real-life reception. That said, any causal relationships among them should not be assumed. For instance, highly readable texts may not necessarily generate greater perceptions of trustworthiness (e.g., Macreadie et al. 2025); while recipients may still change their behavior during emergencies even if they do not fully trust a translation, particularly when no alternative sources of information are available (Rajkhowa et al. 2025). Further empirical evidence is needed to understand how these dimensions interact dynamically

with one another, which would also help conceptualize what constitutes a high-quality translation in crisis settings. Although previous studies have focused on one or more of these dimensions, the present model does not assume that any single dimension should be prioritized over the others without empirical justification. In addition, the recipients' prior knowledge about emergencies, their stylistic expectations of warnings and safety instructions and their trust in the authoritative source of messages could also shape how the translated message is received.

A range of human-centered measures have been used to assess quality in relation to one or more of the four dimensions, including

- Relative preference evaluation, in which bilingual recipients compare and rank different versions of translation (of the same start texts). This method can elicit intuitive user responses to translation quality, often prompted via questions like “Which text do you find most satisfactory to read?” and “Which text would you prefer to recommend to your friends and family?” This method can be used to compare how effective different translation workflows are and the quality of different versions of automated crisis translations.
- Direct assessment, in which authentic users exercise a scalar judgement of the translation, on a scale commonly from 0 to 10. The participants may not have access to the start text or alternative translations. This method could be used to measure all four quality parameters as described above.
- Task-specific comprehension questionnaires provide a measure to explicitly test users' understanding of crisis-related information. This could be particularly useful to flag cases in which “official language” fails to meet the linguistic and cultural norms of the affected populations. Similarly, questions can be designed also to focus on actionability, such as “does it clearly specify what to do or avoid?”
- In the same vein, task-based gap-filling tasks also seek to examine the actual use of translation in terms of successful information uptake, but perhaps within a more controlled setting. It requires participants to restore missing information in a reference text based on the available translation.

Alongside these methods, qualitative approaches, including individual surveys and interviews, focus-group discussions and community quality validation workshops remain powerful ways to yield rich insights into user perceptions of trustworthiness and cultural appropriateness.

Empirical studies conducted by Melbourne research team

This section reviews a series of research initiatives undertaken in Melbourne, Australia, between 2020 (during the COVID-19 pandemic) and 2024 (when Anthony Pym started to take on a more honorary role at the University of Melbourne). These projects focused on multilingual communication aimed at engaging Culturally and Linguistically Diverse (CALD) communities in the context of COVID-19 and other emergencies, such as bushfires and floods. Anthony was the “brain” behind all these initiatives. The focus here is not on seeking advice from experienced translators or accreditation authorities on translation quality, but on hearing from the authentic recipients of the translated public message (i.e., members across various CALD communities) about the potential effects of translated material on them. I was first involved in the COVID-19 project during my PhD candidature to assist with data collection and analysis; and from this project, the empirical data have revealed diversified preferences across various communities, offering more nuanced and contextualized insights than accuracy and grammatical correctness on the text level. Later, I was further involved more closely in the research design and research output preparation in the two subsequent emergency communication projects. I share with the team the ethos that a text is not complete until it has been interpreted. In the field of crisis translation, it is particularly important to examine how authentic readers interpret and respond to the message alongside textual analysis.

These initiatives provide a longitudinal empirical account of how translation quality has been reconceptualized in practice. Some shared features of these projects include: 1) “quality” is situated within a broader socio-cultural context as a communicative success, i.e., the prospective audience have timely access to the translated message that is comprehensible, trustworthy and actionable; 2) non-experts are often recognized as epistemic authorities across the projects, i.e., CALD community organizations, leaders and members; 3) the most advanced technologies available at each point, from neural machine translation to ChatGPT-4omini, are tested in the context of multilingual public messaging.

Early pandemic insights (2020-2022, Hajek et al. 2022; Sengupta et al. 2024)

During the COVID-19 pandemic, the Melbourne research team investigated how CALD communities accessed and evaluated the government-issued translation of public health materials (Hajek et al. 2022). This work situates “quality” within the lived experiences of CALD communities, providing some of the first evidence of community-led and machine-assisted translation practices in crisis settings.

More specifically, interviews were conducted with 58 community members and leaders from African, Pasifika, Afghan, Myanmar, Indian

subcontinent, Arabic and Chinese communities that were resident in Victoria, Australia. Our participants, although being heterogeneous in terms of knowledge expertise, all held a clearly defined role in the community, e.g. represent a community organization, advocate for the community and/or being widely perceived as important people with high standing in their community. Through these interviews, we explore how participants and their communities have followed and received information about COVID-safe behavior, testing and vaccination; we also asked about their level of satisfaction and to what extent they trust the government-issued materials. These questions were then followed by a series of quality evaluation tasks, in which community feedback was collected on authentic materials, e.g., official pamphlets and posters on “How to Handwash” and “After Vaccination” available in public spaces and online.

Here, quality is explicitly defined in terms of whether multilingual healthcare messaging enables users to access, understand, trust and act upon behavior-change messaging in a timely manner (Hajek et al. 2022; Sengupta et al. 2024). It is evaluated by community participants in terms of ease of access, frequency of updates, message completeness and relevance. Quality is signaled by whether translated messages enable recipients to engage effectively with information, especially messages related to vaccination.

For context, as we know, Melbourne was under a nearly nine-month lockdown, making it one of the cities that experienced the longest COVID-19 lockdowns worldwide. In the later stages of the pandemic, the government introduced policies that associated vaccination rates with the easing of lockdown restrictions. Part of the reason lies in the various conspiracy theories and misinformation related to vaccination. In addition to concerns about potential health impacts, vaccination is viewed negatively in some communities. These beliefs include claims that vaccines contain microchips, vaccines alter DNA, and lockdowns and vaccination campaigns were simply implemented as mechanisms of government control. Consequently, the government made a significant effort in seeking effective ways to engage (non-)English-speaking communities for vaccination. During the interviews, most participants whose community relied on translation for information access reported that official translations were often delayed, linguistically inconsistent (mixing two languages), or culturally misaligned (Hajek et al. 2022). For instance, many community participants flagged the official translation of medical information about vaccines and antibodies as overly technical and thus difficult for them and for their people to follow, as these communities were mainly refugees and had relatively lower educational and literacy levels. In other cases, participants complained that the steps described in the official translation for booking vaccination appointments, receiving the vaccine and/or staying safe afterwards were so confusing that they did not know what to do. Some images and colors on the official pamphlets were also perceived as culturally inappropriate or offensive. For instance, several

Pacific communities flagged the issue that using real people's photos in the vaccination jab pamphlet may scare their communities and that cartoons should be used instead; red and black were reported as inappropriate by a few African communities, as these colors may cause negative psychological impacts. As a result, bilingual community leaders and organizations began producing timely, culturally appropriate materials for local dissemination via social media platforms such as Twitter and WhatsApp groups; Google Translate was also reported to be used in some cases.

At the same time, "quality" is also found to be intertwined with trust, which depends not only on textual parameters, i.e., grammatical correctness and idiomatic expression, but also on the perceived credibility of sources (in our case, Victorian Government) and mediators, e.g., NAATI-accredited translators or bilingual community members. Interestingly, we also observed several trust-based strategies adopted by the government in the later stages to boost vaccination rates. These strategies re-positioned mediators and sources of information as originating from within the community itself, including efforts to engage bicultural healthcare workers to produce tailored vaccination promotion videos for their own LOTE (languages other than English)-community and to mobilize community or religious leaders to disseminate vaccine-related information and encourage uptake, via community engagement and outreach initiatives to promote vaccination awareness and public-health messaging .

Our study also argues for space for legitimate "transcreation" (Sengupta et al. 2024), understood as the adaptation of messages to specific audiences, including additions, omissions, reformatting and changes in layout and visual design, all aimed at enhancing the effects of the translated message on its authentic recipients. That said, whether transcreation or literal translation works better should still be tested empirically with authentic community participants. Quality here emerges as relational and context-dependent and in many cases should be co-constructed with its authentic recipient communities.

Machine translation for emergency messaging (2023-2024, Hajek et al. 2024)

The urgency embedded in emergency communication that requires timely responses motivates us to explore the potential of neural machine translation in this domain. Quality is evaluated across a range of machine-assisted workflows and is explicitly conceptualized in relation to "actionability" in our emergency messaging project, i.e., the ability of translated messages to prompt appropriate behavioral responses among recipients (Hajek et al. 2024). Although having "actionability" underlined, urgency and "timeliness" are nevertheless embedded in certain types of emergency communication that requires rapid response, i.e., emergency alerts and crisis broadcasts. Similarly, "understandability" and "trustworthiness" were also tested to understand whether community participants comprehended the intended instructions and

whether they would recommend the translated message to their family, friends and community.

Adopting a reception-oriented perspective, the study examines three translation conditions: raw machine translation, machine translation post-edited by a human (post-edited) and machine translation of pre-edited source texts (pre-edited). The output of each of these workflows was empirically evaluated with members of Spanish-, Greek-, Chinese- and Dari-speaking communities in Melbourne through individual semi-focused interviews. Participants were also invited to complete comprehension tasks after reading the raw machine translation and to indicate their preferences after reading all three versions.

The results of comprehension tests show that most participants across four language groups were misled by ambiguities and distortions in automated translation. Many participants also flagged counter intuitive machine translations that back translate into English as “park your car next to the fire” or “if your car is on fire, park off the road behind a solid structure” (Hajek et al. 2024: 36), which creates risks to public safety and may have serious, even fatal, consequences. Similarly, when providing preference evaluations, none of the participants chose raw machine translation as the preferred option. Taken together, despite clear speed gains, raw machine translation without “human touch” should not be used for emergency alerts and safety instructions. On the other hand, post-edited and pre-edited versions were consistently viewed as more accessible and easier to understand; participants reported that they were more likely to follow the instructions in these versions and share them with family and community members.

This study further forges links between linguistic features and actionability: clearer, more explicit and syntactically simpler translations are more likely to be perceived as easy-to-understand and thus elicit perceived behavioral change. Pre-editing plays an important role in this regard, as it standardizes texts, resolves ambiguities and simplifies syntactic structures. This also explains why the results showed a slight preference among our participants for pre-edited outputs over post-edited ones. At the same time, post-editing processes that used cultural adaptation seem to enhance the perceived legitimacy and trustworthiness of the message source.

Taken together with time-based evaluation, the findings indicate that raw machine translation is faster than pre-editing (particularly when translating into multiple languages), followed by post-editing. Across the four languages, the Dari translators spent the longest time in post-editing machine translation of the same text. The project formed the basis for the following recommendations: (Hajek et al. 2024): 1) the urgency inherent in emergencies justifies the inclusion of automated translation, however, gains in speed associated with fully automated translation should not be traded off against reliability and overall communicative effectiveness (see examples of counterintuitive automated translation); 2) pre-editing that fixes problems in

the ST may offer a more cost-effective solution than post-editing when texts are to be translated into multiple languages; 3) the time required for post-editing tends to increase for less-electronic-resourced languages, such as Dari in our case. Finally, preparatory work conducted prior to emergencies, such as the development of translation memories, glossaries and pre-translated templates, can be used to override raw automated output in time-sensitive situations.

Cultivating a reception-oriented mindset about quality in the GenAI era (2024-2025, Hao and Pym 2026)

Our recent studies on GenAI-translated emergency messaging introduce an explicit pedagogical move toward cultivating a reception-oriented mindset of quality among our translation students (Hao and Pym 2026; Hao and Wang unpublished). A series of studies invited Chinese trainee and novice translators in Melbourne to evaluate emergency-related translations generated by ChatGPT, from the perspectives of their LOTE-speaking (Languages Other than English) communities. When revising ChatGPT initial translations and prompt-based revisions, students were asked: “Is this message easy to understand?” and “Does it allow someone from your community to act on it immediately?”

Guided by these questions, the studies place students in evaluative situations where they must simulate the situational constraints of prospective end users. Likewise, in think-aloud protocol (TAP) tasks, participants were asked whether they would recommend a ChatGPT translation “to a non-English speaking member of your community,” thereby directly embedding reception into the evaluation task. Our findings suggest that participants consistently prioritized Chinese texts (translated from English) with clear subjects and imperative structures, while rejected syntactically complex or ambiguous constructions on the grounds that they would be difficult to process “when people are under stress” during emergencies. One participant explicitly articulated this reception-oriented stance through their complaints about clumsy structures copied literally from the English text that sought to explain the context and rationale behind the evacuation instruction: “The goal is to stay safe... Just tell me what to do!” (Hao and Wang, unpublished).

Interestingly, this perspective also led students to accept non-literal and trans-creative strategies as suggested by ChatGPT, such as generalization (e.g., rendering “do not light a match” as “do not light fire”) and legitimate omissions or additions when these better aligned with the intended behavioral outcome.

At the same time, we argue that reception-oriented evaluation must be paired with a “vigilant, low-trust” attitude (Hao and Pym 2026) towards AI outputs. The classroom experiment shows that over-reliance on GenAI leads students to accept fluent but deeply flawed translations, whereas a vigilant

stance enables them to detect problems and further iterate automated revisions (Hao and Pym 2026). Similarly, participants in the TAP study consistently rejected ChatGPT revisions that introduced omissions which could lead to potentially inappropriate life-saving actions, even when these revisions were stylistically concise and polished (Hao and Wang unpublished). The results of our classroom experiment could possibly have broader implications in (non-)professional crisis translation. Given that this vigilance becomes more than ever required of human in the loop of iterative GenAI workflows, where translation is no longer linear but recursive, involving multiple rounds of prompting, evaluation and/or revision. Translators must therefore continuously decide when to trust, refine, or override AI outputs.

In doing so, students begin to internalize a key principle underpinning our projects: quality is not equivalent to fidelity, but to communicative effectiveness under real-world conditions. In this sense, quality in the GenAI era should no longer be associated merely with a static textual property but with an ongoing, situated and human-centered judgement.

Concluding remarks

This position paper sought to discuss what constitutes a “good” translation in multilingual communication during health crises and disasters. It reviews a longitudinal set of empirical studies conducted in Melbourne, in which “quality” is treated less as an intrinsic property of a translated text and more as related to its effects on prospective recipients under real-world constraints, especially in high-stakes and culturally-loaded contexts. That is, rather than having experienced professionals and accreditation authorities examine the quality of translation against the ST and target-language conventions, quality is ontologically perceived as the degree to which intended audiences can access messages in time (“timeliness”), make sense of them (“understandability”), trust them as credible and relevant (“trustworthiness”) and act on them appropriately (“actionability”). This paper then advocates that, in addition to mono- or bilingual analysis on the text level, quality should also be grounded in user experience and communicative outcomes during crises and emergencies.

A good translation is more than just a correct one, in the sense of accurately reproducing the original meaning and being grammatically error-free and stylistically appropriate. The COVID-19 project demonstrates that community-perceived cultural appropriateness and trust in information sources can substantially influence responses to translated messages. As we have seen, delayed access to official translations that are culturally misaligned and difficult to read may undermine comprehension and behavioral changes, whereas locally mediated translations can enhance community relevance and perceived trustworthiness.

At the same time, machine translation and generative AI systems are reported to have great potential in crisis communication (i.e., speed gains), however, fully automated output that introduces ambiguities and distortions should be avoided as even a small probability of unsafe interpretation may have serious consequences. These increasingly iterative workflows, involving multiple rounds of human-AI interaction, will inevitably require human translators to remain engaged in the loop to exercise situated judgement. More specifically, in crisis settings, humans in the loop could be essential to enhance accountability, recipients' willingness to act upon and share messages within their communities. Achieving this requires a vigilant and low-trust stance towards AI-generated outputs, alongside a deep understanding of the preferences and lived experiences of prospective recipients, including how they process and act upon translated instructions under stress during emergencies.

By repositioning "quality" as relational and situated, this paper calls for quality assessment research that combines textual analysis with user-centered, empirically grounded protocols. Future studies, that said, should test the proposed four dimensions of translation effects across different languages, modalities and institutional settings, to connect reception measures more directly to observable behavioral outcomes.

Declaration of AI use

In preparing this manuscript, the author used ChatGPT5.5 for grammar check. All AI-assisted content was reviewed, edited and verified by the author, who takes full responsibility for the accuracy and integrity of the final text.

References

- Alexander, David E., and Gianluca Pescaroli. 2020. "The role of translators and interpreters in cascading crises and disasters: Towards a framework for confronting the challenges". *Disaster Prevention and Management: An International Journal* 29(2): 144–156.
- Briva-Iglesias, Vincent, and Sharon O'Brien. 2026. "Human-Centered AI Language Technology (HCAILT): An empathetic design framework for reliable, safe and trustworthy multilingual communication". *International Journal of Human-Computer Interaction*: 1–15.
- Cadwell, Patrick, Sharon O'Brien, and Erin DeLuca. 2019. "More than tweets: A critical reflection on developing and testing crisis machine translation technology". *Translation Spaces* 8(2): 300–333.
- Chesterman, Andrew. 1998. "Causes, translations, effects". *Target: International Journal of Translation Studies* 10(2): 201–230.

- do Carmo, Félix. 2025. *Evidence Review Report on the Use of AI for Multilingual Communication in Public Services: With a Specific Focus on the NHS*. Guildford: Centre for Translation Studies, University of Surrey. <https://doi.org/10.15126/901630>.
- Federici, Federico, Minako O'Hagan, Sharon O'Brien, and Patrick Cadwell. 2019. "Crisis translation training challenges arising from new contexts of translation". *Cultus* 12: 246–279.
- Federici, Federico. 2020. "Translation in contexts of crisis". In E. Bielsa and D. Kapsaskis (eds) *The Routledge handbook of translation and globalization*. London and New York: Routledge. 176-189.
- Federici, Federico. 2022. "Translating hazards: Multilingual concerns in risk and emergency communication". *The Translator* 28(4): 375–398.
- Federici, Federico, and Patrick Cadwell. 2018. "Training citizen translators: Design and delivery of bespoke training on the fundamentals of translation for New Zealand Red Cross". *Translation Spaces* 7(1): 20–43.
- Guo, Jing. 2026. *An Empirical Study on Foreign Residents' Expectations of Crisis Translation in China during COVID-19: User-centred Translation Quality as a Multidimensional Quality Metrics Error Typology*. Doctoral thesis. Auckland: The University of Auckland.
- Hajek, John, Maria Karidakis, Riccardo Amorati, Yu Hao, Medha Sengupta, Anthony Pym, and Rita Woodward-Kron. 2022. *Understanding the Experiences and Communication Needs of Culturally and Linguistically Diverse Communities during the COVID-19 Pandemic*. Melbourne: The University of Melbourne. Available at: https://arts.unimelb.edu.au/__data/assets/pdf_file/0020/4071170/Report_CALD_COVID.pdf
- Hajek, John, Anthony Pym, Yu Hao, Maria Karidakis, Ambrin Hasnain, Anila Hasnain, Juerong Qiu, Ke Hu, and Rachel Macreadie. 2024. *Understanding and Improving Machine Translations for Emergency Communications*. Melbourne: The University of Melbourne. DOI: 10.17613/jthe-m639.
- Hao, Yu, and Anthony Pym. forthcoming in 2026. "Can generative AI help students revise translations? Yes, with low, vigilant trust". In I. Feinauer and A. Marais (eds) *The Reality of Revision*. Amsterdam: John Benjamins.
- Hao, Yu, and Yizhou Wang. Unpublished manuscript. "Evaluating GenAI translation for emergency communication: A multimodal think-aloud protocol study".
- Holz-Mänttari, Justa. 1984. *Translatorisches Handeln: Theorie und Methode*. Helsinki: Academia Scientiarum Fennica.
- Macreadie, Rachel, Kadija Bouyzourn, Anthony Pym, and Reine Meylaerts. 2025. "Thick and thin trust in translated cross-cultural vaccination messaging". *Translation Studies* 18(2): 265–282.

- Nida, Eugene. 1964. *Toward a Science of Translating, with Special Reference to Principles and Procedures involved in Bible Translating*. Leiden: E. J. Brill.
- Nord, Christiane. 2005. *Text Analysis in Translation: Theory, Methodology, and Didactic Application of a Model for Translation-oriented Text Analysis*. Amsterdam and Atlanta: Rodopi.
- O'Brien, Sharon. 2011. "Towards predicting post-editing productivity". *Machine Translation* 25(3): 197–215.
- O'Brien, Sharon. 2022. "Crisis translation: A snapshot in time". *InContext* 2(1): 84–108.
- O'Brien, Sharon, and Patrick Cadwell. 2017. "Translation facilitates comprehension of health-related crisis information: Kenya as an example". *Journal of Specialised Translation* 28: 23–51.
- O'Brien, Sharon, and Federico Federici. 2020. "Crisis translation: Considering language needs in multilingual disaster settings". *Disaster Prevention and Management: An International Journal* 29(2): 129–143.
- O'Brien, Sharon, Federico M. Federici, Patrick Cadwell, Jim Marlowe, and Benjamin Gerber. 2018. "Language translation during disaster: A comparative analysis of five national approaches". *International Journal of Disaster Risk Reduction* 31: 627–636.
- O'Brien, Sharon, Michel Simard, and Marie-Josée Goulet. 2018. "Machine translation and self-post-editing for academic writing support: Quality explorations". In J. Moorkens, S. Castilho, F. Gaspari, and S. Doherty (eds) *Translation Quality Assessment: From Principles to Practice*. Cham: Springer International Publishing. 237–262.
- O'Mara, Brigid, and Gemma Carey. 2019. "Do multilingual androids dream of a better life in Australia? Effectiveness of information technology for government translation to support refugees and migrants in Australia". *Australian Journal of Public Administration* 78(3): 449–471.
- Ogie, Rodney I., and Pascal J. Perez. 2020. "Collaborative translation of emergency messages (Co-TEM): An Australian case study". *International Journal of Disaster Risk Reduction* 50: 101920.
- Plitt, Mirko, and François Masselot. 2010. "A productivity test of statistical machine translation post-editing in a typical localisation context". *Prague Bulletin of Mathematical Linguistics* 93(1): 7–16.
- Pym, Anthony. 2023. *Exploring Translation Theories*. London and New York: Routledge.
- Pym, Anthony, Nune Ayvazyan, and Jennifer M. Prioleau. 2022. "Should raw machine translation be used for public-health information? Suggestions for a multilingual communication policy in Catalonia". *Journal of Language Rights & Minorities* 1(1–2): 71–99.
- Reiss, Katharina, and Hans J. Vermeer. 1984. *Grundlegung einer allgemeinen Translationstheorie*. Tübingen: Niemeyer. Trans. Christiane Nord as

- Towards a General Theory of Translational Action: Skopos Theory Explained*. Manchester: St. Jerome, 2013.
- Rossetti, Alessandra. 2019. "Intralingual translation and cascading crises: Evaluating the impact of semi-automation on the readability and comprehensibility of health content". In F. Federici and S. O'Brien (eds) *Translation in Cascading Crises*. London: Routledge. 219–240.
- Seale, Holly, Ben Harris-Roxas, Anthony E. Heywood, Ibrahim Abdi, Abdulmajeed Mahimbo, Lisa Woodland, and Elizabeth Waller. 2024. "'It's no use saying it in English': A qualitative study exploring community leaders' perceptions of the challenges and opportunities with translating and interpreting COVID-19 related public health messaging to reach ethnic minorities in Australia". *PLOS One* 19(2): e0284000.
- Sengupta, Medha, Anthony Pym, Yu Hao, John Hajek, Maria Karidakis, Rita Woodward-Kron, and Riccardo Amorati. 2024. "On the transcreation, format and actionability of healthcare translations". *Translation & Interpreting* 16(1): 121–141.
- Staiano, Maria Carmen, Lifeng Han, Johanna Monti, and Francesca Chiusaroli. 2025. "ITALERT: Assessing the quality of LLMs and NMT in translating Italian emergency response text". In *Proceedings of Machine Translation Summit XX: Volume 1*. 566-577.
- Todorova, Miglena. 2024. "Translation and emergency". In C. V. Angelelli (ed) *The Routledge Handbook of Translation and Sociology*. London and New York: Routledge. 383-399.
- Toury, Gideon. 1980. *In Search of a Theory of Translation*. Tel Aviv: Porter Institute for Poetics and Semiotics.
- Toury, Gideon. 1995/2012. *Descriptive Translation Studies – and Beyond*. Revised edition. Amsterdam and Philadelphia: John Benjamins.
- Vermeer, Hans J. 1989. *Skopos und Translationsauftrag*. Heidelberg: Institut für Übersetzen und Dolmetschen.

How does multimodal creativity travel? A reception perspective on cyberdubbing

JUERONG QIU

University of Melbourne, Australia

This article examines three cyberdubbed versions of a trailer of Zootopia 2 on Bilibili to explore how multimodal creativity is produced and received in participatory media environments. It shows how cyberdubbers combine dialogue rewriting, vocal performance, subtitles, visual design and editing to create humor, authenticity, local resonance and media memory. Audience responses in danmu and forum comments further show that these creative choices are not simply received but cooperatively interpreted, explained and extended. This cooperative interpretation of authenticity provides the basis for contingent audience trust in cyberdubbers as alternative audiovisual translation products.

Keywords: cyberdubbing; audience reception; creativity; authenticity; cooperation

Introduction

Between 1949 and 1994, when foreign films entered Chinese mainland, they were predominantly distributed through dubbing. This was largely due to the low literacy rate of the population and the influence of Soviet film practices (Du 2018). As a result, dubbing became institutionalized within a state-controlled film production system. For example, the Shanghai Film Dubbing Studio became the primary institution responsible for translating and dubbing foreign films into Mandarin (Feng 2025). In this context, foreign films were translated to ensure immediate comprehensibility and lip synchronization and then performed by professional voice actors in ways that aligned with local audiences and ideological expectations (Zhang 2004).

Dubbed films in this period, commonly referred to as “译制片”, are distinct from contemporary dubbing practices in China due to their linguistic style, which Du (2018) describes as “Chinese dubbese.” This style is marked by phonetic idiosyncrasies, including drawls, prestos, padding, and unnatural pauses rarely found in everyday Mandarin. Although this vocal style has

gradually faded in contemporary dubbing, Chinese dubbese has recently re-emerged as a discursive resource on social media platforms. No longer confined to the screen, it is imitated, parodied, and remixed by Internet users in videos, comments and memes (Chen 2025). In this sense, a once institutionalized translation practice has been recontextualized within participatory digital culture.

In recent years, foreign films with Chinese subtitles show significantly higher viewing rates among viewers aged between 18 and 30, while for family-oriented viewing contexts, dubbed versions have a higher proportion of viewership, particularly on television and projector-based platforms (Cai, 2026). In dubbing practices, productions have experienced a deprofessionalizing process as dubbed films are no longer performed exclusively by professional voice actors. Online celebrities and idol actors are increasingly involved in dubbing, often to attract younger audiences. Further, advances in digital editing technologies and audio production tools have allowed viewers to produce derivative works through remixing, modification, revoicing, etc. One such practice is cyberdubbing, defined as dubbing activities produced in non-professional and non-commercial contexts and distributed primarily through online platforms (Baños and Díaz-Cintas, 2024). Although cyberdubbing takes diverse forms and is shaped by different motivations, including altruistic, fan-driven and activist purposes, its reception remains underexplored.

The case of *Zootopia 2* offers an opportunity to examine how cyberdubbed productions are created, compared with official dubbing and collaboratively interpreted by audiences in online spaces. The first *Zootopia* (2016) achieved remarkable success in China, grossing \$236 million and becoming the highest-grossing Hollywood animated film in the market (Masunaga 2025). This success created high expectations for the sequel after a nine-year interval. However, the announcement of the Chinese-language voice cast for *Zootopia 2* triggered widespread debate, particularly regarding the increasing reliance on celebrity voice actors. Many viewers questioned whether such casting decisions compromised performance quality and professional standards. Criticism intensified following promotional events and the premiere, where audiences noted issues such as unclear pronunciation, mismatched vocal performance and perceived lack of professionalism (Zhu 2025). Following its premiere on 26 November 2025, cyberdubbed versions of *Zootopia 2* trailers quickly circulated on user-generated video platforms, such as Bilibili and Xiaohongshu. These cyberdubbings offer sites where audiences and creators evaluate, reinterpret and at times challenge official dubbing practices.

Thus, the present study examines cyberdubbed trailers of *Zootopia 2* on Bilibili, focusing on three highly viewed videos and their accompanying danmu and forum comments. Through a multimodal analysis of selected cyberdubbed videos, together with keyword and attitude analysis of user-

generated content, this study investigates how cyberdubbers use different multimodal elements to produce creative alternatives to official dubbing, and how audiences respond to, extend and evaluate these creative practices. Building on Pym's (2020) insight that new forms of translation, such as adaptation, rewriting and transcreation, reshape the notion of authenticity, the study further examines how successful audience cooperation contributes to the formation of trust in cyberdubbers.

Multimodality and cyberdubbing

Audiovisual translation is inherently multimodal, as meaning is created and conveyed through the auditory-verbal channel, the auditory-non-verbal channel, the visual-verbal channel and the visual-non-verbal channel (Gottlieb 1998). Kress and van Leeuwen (2001) define multimodality as the use of several semiotic modes and the ways in which these modes are combined within a communicative product or event. Chaume's (2012) analytical framework provides a specific way to examine the multimodal structure of audiovisual texts. He conceptualizes audiovisual texts as being composed of meaning codes transmitted through the acoustic and visual channels. These codes interact with one another to shape meaning, interpretation and cinematography.

Chaume's framework has been adapted to the study of different audiovisual translation modes. Romero-Muñoz (2023) uses Chaume's classification of meaning codes to examine how objectivity and subjectivity coexist in audio descriptions of Netflix Spanish drama and thriller series. In the research on subtitling for the D/deaf and hard-of-hearing (SDH), López Salvador and Spiteri Miggiani (2025) use Chaume's taxonomy to compare conventional and creative SDH, showing how creative freedom reshapes the descriptive quality of SDH tags and makes subtitlers' authorial input more visible.

This framework can also be applied to dubbing, since dubbing involves replacing the original dialogue track with a target-language recording that reproduces the original message while synchronizing the target-language voice with the actors' lip movements (Díaz-Cintas and Orero 2010). Cyberdubbing in Chinese mediascape further complicates this. Unlike the contexts in which dubbing and subtitling are conventionally treated as separate audiovisual translation modes, Chinese dubbed products, including both official dubbing and cyberdubbing, are commonly accompanied by Chinese subtitles. Subtitles also form part of its multimodal presentation and need to be included in the analysis.

On Bilibili, this multimodal structure is further extended by danmu comments. Massidda (2025) identifies danmu as a form of in- and out-show comment within creative audiovisual translation in the digital age, alongside

in-show titles and web titles. Danmu, or danmaku, in Japanese, refers to a commenting system in which viewers post timeline-synchronized comments that appear across the screen while audiovisual content is playing, creating a “watching together” environment on Asian video-sharing and streaming platforms such as Bilibili, iQIYI, Tencent, Youku and Nico Nico Douga (Massidda, 2025). Because danmu comments are visually overlaid onto the video during viewing, they add another visual-verbal layer to the audiovisual text and turn audience responses into part of the audiovisual encounter. Cyberdubbing on Bilibili therefore involves not only the interaction between voice, image and subtitles, but also the participatory layer of audience comments, making it a particularly rich site for multimodal analysis. This study therefore examines cyberdubbing as a multimodal and participatory audiovisual translation practice, showing how cyberdubbers make use of acoustic and visual meaning codes to produce creative reinterpretations and how these creative strategies are received and evaluated by audiences.

Methods

This case study uses Chaume’s (2012) analytical framework to analyze the creative multimodal practices found in cyberdubbing. The analysis focuses on how cyberdubbers employ voice acting styles, regional dialects, humorous reinterpretations of dialogue and experimental subtitle practices to convey meaning. Danmu and forum comments are then analyzed to examine how audiences engage with, extend and evaluate these creative strategies.

I first conduct a keyword search on Bilibili using the term “疯狂动物城2 配音” (*Zootopia 2* dubbing). The search was performed in an incognito browser window without logging into a user account and without applying additional filters, in order to minimize the influence of personalized recommendations and reflect the platform’s default visibility structure. Of the 1,000 total search results, I arranged them based on relevance and view count and selected the three most viewed cyberdubbed videos based on the same international trailer of *Zootopia 2*. Focusing on the same source material allows for clearer comparison of translation strategies and dubbing choices across different versions. These videos attracted substantial audience engagement. Table 1 provides key information about the selected videos, including upload date, view count and number of comments.

This study examines how cyberdubbers used different multimodal elements in their practice and how the videos received on Bilibili. Video 1, 天津话配音《疯狂动物城2》，听过吗弟弟! (*Zootopia 2* Dubbed in Tianjin Dialect, Have You Heard This, Bro?), posted by 没6的沈龙Loong, is explicitly framed as a non-official, recreational production (“just dubbing for fun” in the video description). It is created by two cyberdubbers with the

female dubber voicing all female characters and the male dubber voicing all male characters. Video 2 (北京话配音《疯狂动物城2》(Zootopia 2 Dubbed in Beijing Dialect)), posted by 画家常小亮, does not provide a description and appears to be an individual production, dubbed entirely by a single male voice. Video 3 (假如用CCTV6老译制片的方式打开《疯狂动物城2》(What If Zootopia 2 Were Dubbed in the Style of Classic CCTV-6 Chinese dubbese?)), posted by 配音演员张小麦, reflects a more organized and professional production context. The uploader acknowledges the support of multiple voice actors in the description and the video is dubbed by a team of eleven professional dubbers, each performing a specific character.

Table 1: Basic information about the three videos

Video number	Time posted	Duration	Views	Likes	Danmu comments	Forum comments
V1	28 Nov 2025	01:33	1,572,000	47,000	339	756
V2	03 Dec 2025	02:08	643,000	22,000	329	630
V3	09 Dec 2025	02:15	571,000	23,000	232	538

The analysis focuses on two types of data. First, this analysis incorporates multiple channels of information, including the visual presentation of subtitles (e.g., speed, font, color and style), visual elements such as inserted image and filters, and the acoustic channel (e.g., pronunciation, accent and intonation). Table 2 shows how selected codes from Chaume's (2012) framework are operationalized in this study. Examples are analyzed in detail in the Results section below.

The second dataset consists of danmu comments and forum comments posted by viewers. Danmu comments are time-synchronized with the video and include both timestamp and content, while forum comments contain the sender's name, posting time and content.

Keyword analysis was conducted to identify the most frequently mentioned terms in the danmu comments. Since the comments were written in Chinese, Jieba, an open-source Chinese word segmentation tool, was first used to segment the comments into individual lexical units. Keyword frequency was calculated by word occurrence rather than by comment. Therefore, repeated occurrences within the same comment, such as repeated instances of “哈哈” (haha), were included in the count (See Table 6). This analysis was used to identify the aspects of cyberdubbing that attracted the most immediate audience attention.

Table 2: Multimodal analysis using Chaume’s framework: Codes and examples

Code	Description	Examples
Linguistic code	How cyberdubbers create comic effects through dialectal vocabulary, humorous rewriting, colloquial expressions, forms of address and Chinese dubbese.	Humorous rewriting of the original trailer dialogue; CCTV-style “Chinese dubbese” in Video 3
Paralinguistic code	Vocal features that shape the delivery of the dubbed dialogue, including pronunciation, accent, intonation, rhythm, pitch, emotional tone and voice quality.	Tianjin or Beijing accent; playful or comic intonation; added lines in response to non-subtitled whispers, gasps or exclamations in the original trailer.
Iconographic code	Visual symbols, images, objects and culturally recognizable signs that contribute to meaning making.	Fox and rabbit ears worn by the dubbers in Video 1; rabbit and fox drawings in Video 2; CCTV-6 logo in Video 3;
Photographic code	Visual qualities of the image, including color, lighting, contrast, brightness, filters and overall image texture.	Faded or blurred color filter in Video 3 to imitate older dubbed films; brighter and more playful visual presentation in Videos 1 and 2.
Mobility code	Coordination between dubbed dialogue and visual movement, including lip synchronization, facial expressions, gestures and action timing.	Target utterances timed to match characters’ mouth openings and closings; humorous lines delivered in response to characters’ gestures or facial expressions
Graphic code	Written language and graphic text appearing on screen, including subtitles, titles, captions, danmu comments, font, size, color, placement and subtitle style.	Stylized subtitles in Video 1; large subtitles and decorative layout in Video 2; conventional white subtitles with black outline in Video 3; danmu comments overlaid onto the video screen.
Editing code	The organization and rearrangement of visual elements.	Dubbers’ performance window inserted in Video 1; original trailer frame resized to create space for subtitles and graphic elements in Video 2; reconstructed CCTV-6-style layout in Video 3.

Note. The remaining four codes are excluded because the associated elements are not applicable to the selected clips.

Attitude analysis was conducted through inductive coding of danmu and forum comments. Rather than using a pre-existing sentiment dictionary, the attitude codes were developed from recurring evaluative expressions in the comments. Comments containing explicit affective or evaluative wording, such as “太好了” (so good), “好笑” (funny), “地道” (authentic), “很烦” (annoying), “尴尬” (awkward), or “听不懂” (difficult to understand), were identified and grouped according to their evaluative focus. These codes were then organized into specific categories, such as performance quality, local authenticity, stylistic recognition, comprehension, etc. (See Table 6). In this

sense, attitude analysis captured the evaluative orientation of the audience comments, while thematic analysis was used to develop and interpret the categories.

Results

Creativity in cyberdubbings

Unlike conventional dubbing, which typically aims to minimize visual distraction and maintain lip synchronization, cyberdubbing foregrounds its own production process and creative intervention. In the three videos analyzed in this study, creativity is expressed not only through translation choices but also through various visual codes.

In Video 1, the uploaders use a creative font, a rounded typeface that produces a softer and more playful visual effect (see Figure 1). The subtitles have a blue glow and white outline, making them visually prominent against the video background. The video title appears in the upper left corner of the screen, while a small window is placed in the lower left corner showing the two cyberdubbers performing the voices in real time. The male dubber wears fox ears and voices all male characters, while the female dubber wears rabbit ears and performs all female characters. These graphic and editing codes not only reveal the dubbing process but also playfully match the performers with the animal characters in the film.

In Video 2, the original trailer is reduced in size to create additional space for subtitles and graphic elements (see Figure 2). The subtitles are larger than standard subtitle size and are presented in the Yahei font, which is commonly used in digital media for its clarity and readability. On both sides of the subtitles, simple line drawings of a rabbit and a fox are added as decorative elements. The uploader's username appears in the upper right corner, while the video title is placed at the top of the frame. These visual codes give the video a layout that resembles a fan edited media product rather than a conventional subtitled or dubbed video.

Video 3 imitates the esthetic conventions of classic official dubbed films in Chinese Mainland. As shown in Figure 3, the video carefully recreates many of the stylistic details associated with foreign films broadcast on China Central Television Channel 6 (CCTV-6). For instance, a CCTV-6 logo is displayed in the upper left corner of the screen. On the left side of the frame, a vertical text presents information about the film title and the country of origin, a format commonly used in 1990s' dubbed films. The subtitles also follow the conventional style of traditional dubbing subtitles, using the standard Yahei font with white characters and black outlines. The color of the video appears noticeably blurred and faded compared with Videos 1 and 2. This visual treatment creates a nostalgic atmosphere reminiscent of older films.

Figure 1: Screenshot of Video 1 at 00:14**Figure 2:** Screenshot of Video 2 at 00:14**Figure 3:** Screenshot of Video 3 at 00:14

Video 3 imitates the esthetic conventions of classic official dubbed films in Chinese Mainland. As shown in Figure 3, the video carefully recreates many of the stylistic details associated with foreign films broadcast on China Central Television Channel 6 (CCTV-6). For instance, a CCTV-6 logo is displayed in the upper left corner of the screen. On the left side of the frame, a vertical text presents information about the film title and the country of origin, a format commonly used in 1990s' dubbed films. The subtitles also follow the conventional style of traditional dubbing subtitles, using the standard Yahei font with white characters and black outlines. The color of the video appears noticeably blurred and faded compared with Videos 1 and 2. This visual treatment creates a nostalgic atmosphere reminiscent of older films.

Table 3: Example of subtitles mismatch

Start time	Original line	Dubbed version	Subtitled version
00:02	Hopps and Wilde in pursuit of suspect.	朱迪和尼克在此 嫌疑犯别想逃	朱迪和尼克 正在追捕嫌疑人
00:04	Back off, rookies.	这轮不到你插手	让开 菜鸟
00:05	We got 'em! Rolling spikes!	交给我们就行 伙计们钉刺准备	我们要抓住他们了 放钉刺
00:07	Rolling spikes!	钉刺给我扎	放钉刺

Unlike Videos 1 and 2, in which the subtitles correspond exactly to the dubbed dialogue, Video 3 deliberately reproduces a mismatch between written subtitles and spoken dialogue. In classic dubbed films, subtitles were present

but did not fully correspond to the dubbed conversation. Minor discrepancies between what viewers read and what they heard were therefore common. During the dubbing process, the script often needs to be adjusted on the spot to match lip movements and other performance requirements without coordinating with the subtitle team (Zhang 2004). This characteristic is deliberately reproduced in Video 3. As shown in Table 3, in order to match the lip movements and syllabic rhythm of the English dialogue, the actual dubbed version is often longer than the corresponding subtitle version. Through this mobility code, the video not only parodies the historical style of Chinese official dubbing in the 20th century but also demonstrates a form of intertextual creativity that draws on the audiences' shared cultural memory of the tradition. These choices are recognized, praised and explained by viewers on danmu

00:00.9: 经典配音对不上字幕 [Classic dubbing where the subtitles don't match the spoken lines.]
 00:04.0: 细节影片后面加上国家地区名 [Nice detail: the country/region name is added after the film title.]
 00:17.7: 貌似上个世纪的产物 [Feels like something from the last century.]

and forum comments.

废号*: 为什么老电影喜欢字幕跟配音不一样? [Why do old films often have subtitles that don't match the dubbing?]
 春风* replies: 因为字幕是根据原台词在胶片上制作的而配音为了对上口型要对台词进行改变。[Because the subtitles are created based on the original script and embedded onto the film, while the dubbed lines are adjusted to match lip movements.]

Table 4: Basic information about subtitles in each video

	Subtitle lines	Lines omitted	Lines added	Max length	cps
V1	41	15	6	15	8
V2	52	5	4	12	12
V3	58	0	0	16	8

Further, Table 4 shows that Video 3 does not add or omit any original subtitles, while Videos 1 and 2 exhibit a substantial degree of omission and addition. The omissions in Video 1 are mainly due to the uploaders deleting certain scenes from the original video. In Video 2, the omissions primarily involve brief discourse markers or interjections, such as “ow,” “hi,” and “oops” The additions, on the other hand, mainly take the form of extra monologues or inserted lines that enhance the dramatic effect of the scene, often contributing to a more humorous and exaggerated tone.

For example, in Video 1, during Judy Hopps and Nick Wilde's pursuit of Gary De'Snake, Gary reaches the tube station. Nibbles Maplestick warns them not to follow him into the Red Line. Gary jumps into the Red Line and

Judy immediately follows. Then, Nick reluctantly joins in after her. In this moment, an additional line is inserted: “你知道跳水大爷吗” (Nick: Do you know those old guys who go diving?). This line refers to a well-known local phenomenon in Tianjin, where middle-aged and elderly men perform highly distinctive public diving. Before jumping, they often deliver humorous opening remarks and interact playfully with the audience, creating a lively atmosphere filled with cheers and exclamations. By invoking this culturally specific reference, the cyberdub reframes the act of jumping into the Red Line as a performative and humorous spectacle, enhancing both the local flavor and comedic effect of the scene.

In Video 2, two hippo police officers are covered in spikes after being accidentally attacked by their porcupine colleague. As they attempt to squeeze through the doorway into a meeting room, the situation becomes highly comical. Additional two off-screen lines are inserted to amplify the humor: “哎呦喂穿孔打钉” (Ouch, looks like they’ve been studded full of piercings!) and “成亚比了嘿” (Looks like they’ve turned into some kind of punk!).

Further, creativity is also manifested in the highly localized dubbing practices adopted in the videos. Two of the analyzed versions are dubbed in Tianjin dialect and Beijing dialect respectively, incorporating distinctive local expressions, intonation patterns and dialectal vocabulary. The notion of “local” here refers to a strong regional flavor that resonates with viewers familiar with these speech styles. At the same time, the dialectal performances produce an effect that is both “exotic” but accessible for broader audiences. For viewers who are less familiar with the dialects, danmu comments often provide additional explanations or humorous interpretations, helping others follow the dialogue while reinforcing the participatory nature of the viewing experience.

In this scene, Nick and Judy urgently need to get to another part of the city as part of their investigation. Nick asks Flash, a sloth known for his extremely slow movements and speech, for help.

Original line: We need to get across town.
 Video 1: 天津站到瓷房子50块钱走不走 [Tianjin Railway Station to the Porcelain House, 50 yuan, deal or not?]
 01:18.4: 天津站到瓷房子, 三号线地铁坐到营口道下, 骑小黄五分钟骑到。花不到四块钱, 你要给我50? [From Tianjin Railway Station to the Porcelain House, you can take Line 3 and get off at Ying Kou Dao. You then use shared bike. It takes less than four yuan. You want to give me 50 yuan?]
 01:19.0: 用的了这么多吗? [Do you need to pay that much?]
 01:20.3: 《50块钱》 [FIFTY DOLLARS]

The Tianjin dialect version demonstrates creativity through strong localization and cultural referencing. In the original line “We need to get

across town” the cyberdubbers reinterpret the situation as a taxi negotiation in Tianjin. The line replaces the generic spatial reference “across town” with two well-known local landmarks. Nick’s proposal of a 50-yuan fare, presented as excessively high, adds humor while signaling the urgency. This reinterpretation evokes everyday street interactions characteristic of Tianjin’s local speech culture. Danmu comments further reinforce this effect, as viewers highlight the absurdity of the price and engage with the humor. The humor lies in the exaggerated specificity of the route and the familiar negotiation style, which resonates with local audiences while remaining accessible to wider viewers through shared context and participatory commentary.

Video 2 similarly relies on localization but employs a different strategy by embedding the dialogue in Beijing’s spatial and sociolinguistic references.

Original line: Nick: No snake has set foot in Zootopia in forever.

Video 2: 好多好多年啊 都没有长虫 进二环了 [It’s been many, many years since a snake has entered the Second Ring Road.]

00:29.8: 别说二环了, 现在北京城里都没有什么蛇 [Don’t even mention the Second Ring Road. There are hardly any snakes left in Beijing city now.]

00:29.8: 二环笑死哈哈哈哈哈, 合着还得给长虫半个京牌是吗... [The Second Ring Road bit is killing me, hahaha. So are we supposed to give the snake half a Beijing licence plate too?]

00:30.8: 进二环太地道了哈哈哈哈哈 [“Entering the Second Ring Road” is so authentic, hahaha.]

The reference to Beijing’s Second Ring Road, a well-known geographic and symbolic boundary in the city, within the second ring is considered the center of the city, replaces the fictional location of Zootopia. The term “长虫,” a colloquial and slightly humorous way of referring to snakes in northern dialects (not just limited to Beijing dialect), further reinforces the regional tone. This localization creates humor by framing the snake as if it were a migrant attempting to enter Beijing’s central district, echoing broader cultural associations with residency permits and urban access. Danmu comments amplify this interpretation, with viewers joking that the snake would need a Beijing license plate to enter the Second Ring Road. Through this playful reinterpretation, the cyberdub transforms the original narrative line into a culturally specific joke that draws on shared knowledge of the locale.

Another creative strategy is the incorporation of English sounding elements within the Chinese dubbing. Unlike straightforward transliteration, these examples involve a hybrid form that incorporates English phonetic patterns with Chinese expressions.

Original scene: Nibbles, Judy and Nick enter the reptile hangout, a secret speakeasy in Marsh Market.

Video 1: Chow why did~

In Video 1, a new line is introduced in a scene where no dialogue exists in the original version. The utterance “chow why did~”, with the tilde indicating a wavering intonation, mimics the phonetic contour of the Chinese derogatory phrase “臭外地的” (damn outsiders). This example involves both the linguistic and paralinguistic codes. At the linguistic level, the cyberdub creates a pseudo-English phrase that looks like an English utterance but encodes a recognizable Chinese expression through sound. At the paralinguistic level, the wavering intonation reinforces the phonetic resemblance and helps make the joke identifiable to Chinese-speaking viewers. The humor is therefore produced through a playful friction between linguistic appearance and phonetic familiarity.

ST: Quilda (a female African crested porcupine): Rolling spikes!

Video 2: 扎死他 独一无二

A similar strategy can be observed in Video 2. In the original line, Quilda exclaims “Rolling spikes!” which is rendered as “扎死他 独一无二.” The first part, “扎死他,” conveys the action of attacking by rolling spikes, thus maintaining semantic alignment with the source text. The second part, “独一无二,” imitates the sound of the English slogan “just do it.” This phonetic play introduces an additional layer of meaning, suggesting decisiveness and immediacy in Quilda’s action. Here, the translation operates simultaneously at the level of meaning and sound, combining semantic translation with phonetic creativity. This hybrid strategy enhances the expressive force of the line while embedding it within a broader network of recognizable cultural references.

What do danmu and forum comments tell us?

The keyword analysis, complemented by thematic interpretation of danmu and forum comments, shows that the laughter marker (haha) dominates across all three videos, suggesting that audience engagement is primarily affective and performative (Table 5). Further, Table 6 shows a clear dominance of positive audience evaluations across all three videos, with negative responses remaining minimal and marginal. These two sets of data provide clues about how audiences perceive and evaluate the three videos. These evaluative patterns, in turn, offer insight into how trust in cyberdubbing is formed and negotiated.

Table 5: Top 3 most frequent keywords in danmu comments

Video	Keyword	Translation	Occurrences
Video 1	哈哈	Haha (laughter)	176
	天津	Tianjin	33
	配音	Dubbing	15
Video 2	哈哈	Haha (laughter)	145
	二环	The 2nd Ring Road	27

	北京	Beijing	25
Video 3	哈哈	Haha (laughter)	99
	译制	Chinese dubbese	36
	字幕	Subtitle	24

Table 6: Audience attitudes towards cyberdubbed videos

Video	Attitude	Category	Codes (frequency)
V1	Positive	Humor and engagement	Hilarious (45), many memes (24), enjoyment (12)
		Performance quality	Good dubbing (15), better than official dubs (12), understandable (4)
Local authenticity		Authentic Tianjin expression (28), local accent (13)	
V2	Negative	Performance issues	Awkward delivery (3), mixed accent (1)
		Comprehension	Difficult to understand (1)
V3	Positive	Humor and engagement	Hilarious (38), enjoyment (16)
		Performance quality	Good dubbing (23), better than official dubs (8), Better than official dubs (6)
		Local authenticity	Authentic Beijing expression (24), authentic accent (13)
V3	Negative	Performance issues	Sounds uncomfortable (1)
		Humor and enjoyment	Fun (4), enjoyment (2), nostalgia (4)
		Stylistic recognition	Excellent Chinese dubbese (25), attention to detail (14)
	Negative	Performance issue	Unpleasant voice quality (1)

In Table 6, negative evaluations across all three videos are limited and largely confined to minor issues of performance or comprehension, while there are recurring codes regarding positive comments across the three videos.

In Videos 1 and 2, localization is a central evaluative dimension, as indicated by the frequent appearance of place-related expressions in danmu. In Video 1, “配音” (dubbing) directly reflects audience attitudes toward the cyberdubs. Of the fifteen occurrences, most express strong approval, such as “配音不错” (nice dubs). At the same time, five comments explicitly compare the cyberdub with the official version, often criticizing the latter and expressing a clear preference for the former. What is salient here is not the evaluation alone, but its comparative nature. These comparisons position cyberdubbers as more credible and appealing alternatives, suggesting that preference is not pre-given but relationally constructed through contrast with what is perceived as inadequate official dubbing.

In Video 2, the evaluation focuses on authenticity. Here, authenticity does not refer to staying true to the source text, but to the perceived credibility of the target-locale performance. Expressions such as “正宗老北京” (real old Beijing) and “地道北京味儿” (authentic Beijing flavor) suggest that

viewers evaluate the dubbing in terms of how convincingly it performs “Beijingness.” The frequent occurrence of “二环” (the Second Ring Road) further anchors this response in shared knowledge of the locale. Humor is not only recognized but also extended, as viewers build on the “二环” joke in danmu. In this sense, the credibility of the cyberdubbing depends less on its closeness to the source text than on its ability to b target-locale linguistic and cultural resources.

In Video 3, keywords such as “译制”, “字幕” and “配音” indicate that viewers are responding less to content than to stylistic reconstruction. Evaluation focuses on whether the video successfully reproduces the conventions of “译制片” (Chinese dubbese). Codes such as “excellent Chinese dubbese” and “attention to detail” indicate an appreciation of stylistic imitation. Here, legitimacy is tied to recognizability: the more convincingly the video evokes a familiar dubbing esthetic and shared cultural memory, the more it is accepted as credible.

Discussion: How does creativity cross over to audiences?

The analysis above has shown that the three cyberdubbed videos mobilize different combinations of meaning codes to demonstrate distinct forms of creativity. These creative practices are recognized, interpreted and evaluated by viewers through danmu and forum comments. This discussion examines what viewers perceive as authentic and what this authenticity is oriented towards. It then considers how authenticity is constructed through multimodal elements and how such perceptions are later transformed into cooperation and trust.

As mentioned, authenticity is not understood as staying true to the source text. Rather, it refers to the extent to which a cyberdubbed performance is perceived as locally or stylistically convincing within a specific context. As Pym (2000) puts it, authenticity is an effect produced through the multiplication of variations and variations do not appear authentic because of one isolated marker. In Videos 1 and 2, authenticity is most visibly produced through the linguistic and paralinguistic codes. Dialectal vocabulary, colloquial expressions, humorous rewriting and forms of address integrate with accent, intonation, rhythm and vocal delivery to create locally recognizable performances.

At the same time, authenticity does not require every detail to be immediately understood by all viewers. Some culture-specific references, such as “瓷房子” (Porcelain House) and “二环” (the Second Ring Road), may sound recognizably local while still requiring additional knowledge of the locale to understand the jokes. This is where danmu and forum comments become important. Viewers do not merely receive authenticity as a finalized

effect; they help make it visible by explaining, extending and responding to these local details. The esthetics of detail is therefore amplified through audience participation, as viewers collectively identify what sounds authentic, clarify references for others and make more jokes around them.

While Videos 1 and 2 produce authenticity mainly through dialectal vocabulary and regional vocal performance, Video 3 constructs a different kind of authenticity: mimicking Chinese dubbese and reproducing classic dubbed-film esthetics. This authenticity is produced through (1) paralinguistic features such as tone, rhythm, intonation and delivery, (2) linguistic features associated with translated-film dialogue and (3) visual codes such as subtitle style, faded color, CCTV-6-style layout and broadcast-like editing. The video appears authentic because all these elements activate a recognizable media memory of older Chinese dubbed films and the nostalgic feelings attached to that viewing experience.

Cyberdubbers of the three videos draw on these meaning codes that allow to create interpretations and invite audience participation. These codes function as hooks: viewers who recognize them can explain, extend or play with them, while viewers who do not fully understand them can ask questions, read responses and gradually make sense of the performance. This process aligns with Pym's (2021) observation that communication is not always governed by complete transparency or immediate understanding. Difficulty, ambiguity and partial understanding may themselves become part of cooperative communication. In this sense, communicative success becomes a "cooperative outcome of the communication act" (Pym 2021, p. 12).

We can therefore observe effort distributed across cyberdubbers and viewers who left comments. The former gain visibility, engagement and potential recognition within the dubbing field. The latter, in turn, gain humor, cultural insight and participatory experience. The willingness of both parties to invest effort is closely tied to the benefits involved. As Pym (2015) argues, cooperation depends on the distribution of effort and reward. When the perceived gains are sufficient, participants are more willing to engage.

It is through this cooperative recognition of authenticity that trust begins to emerge. When viewers identify the cyberdubbers' choices as locally or stylistically convincing, they are also recognizing the cyberdubbers' competence. This competence is not established by institutional authority but demonstrated through performance and confirmed through audience interaction. Viewers' explanations, jokes, questions and appreciations therefore do more than extend the meaning of the cyberdubbed videos; they also help validate the cyberdubbers as capable and credible players in audiovisual translation.

This helps explain why some audience comments position cyberdubbed versions against official dubbing. Statements such as "this is much better than the official dubbing" or "I would go to the cinema if it were dubbed like this" do not simply express enjoyment. They suggest that these viewers perceive

cyberdubbers as more responsive to audience expectations and more attentive to local and stylistic details than some dubbers in the official version. Trust here is thus formed relationally, through comparison with official dubbing and through the successful cooperative interpretation of cyberdubbed performances.

These viewers' claims should, however, be understood cautiously. Active commenters represent only a minority of viewers, while many users are likely to be lurkers who watch the videos and read comments without posting their own responses. Moreover, enthusiastic comments should not be read as direct evidence of actual behavioral intention. A statement such as "I would go to the cinema if it were dubbed like this" may be exaggerated, performative, or shaped by Bilibili's long-standing culture of playful, locally flavored fan production. Its significance lies less in whether the viewer would literally buy a cinema ticket than in the evaluative orientation it reveals. Such comments indicate that some viewers recognize and appreciate the cyberdubbers' creative choices, especially when these choices appear more locally attuned, more stylistically convincing, or more engaging than official dubbing.

Conclusion

This study uses three cyberdubbed versions of a *Zootopia 2* trailer as a case to examine how creativity in cyberdubbing is produced and evaluated within participatory media environments. Drawing on Chaume's (2012) framework of meaning codes, the study shows that cyberdubbing creativity is not limited to verbal rewriting or vocal performance, but is produced through the interaction of linguistic, paralinguistic, graphic, iconographic, photographic, mobility and editing codes. Through these codes, cyberdubbers create dialectal humor, localized cultural references, stylized dubbese, subtitle-dubbing mismatches and recognizable media memories. These creative strategies do not operate in isolation, but invite audiences to recognize, interpret and extend them.

The findings further show that audience engagement with cyberdubbing is not merely receptive but cooperative. What viewers evaluate is not on whether the cyberdubbed versions stay true to the source text, but whether they appear authentic in the situated context. Authenticity here refers to performances that are perceived as locally or stylistically convincing. Videos 1 and 2 construct authenticity mainly through dialectal expression, local knowledge and paralinguistic performance, while Video 3 produces a different form of authenticity through Chinese dubbese, subtitle-dubbing mismatch, broadcast-like visual design and nostalgic media memory. In this sense, meaning is co-constructed between cyberdubbers and a subset of active viewers, with danmu and forum comments functioning as an interpretive layer that supports and redistributes communicative effort. As audiences recognize

cyberdubbers' performances as locally or stylistically convincing, they also begin to attribute preference and credibility to these cyberdubbers. Trust is therefore not an additional theme separate from creativity and reception, but emerges from the successful cooperative interpretation of authenticity.

These findings also have implications for professional audiovisual translation practice. When official dubbing does not meet audience expectations, it can be readily challenged by alternative forms of production outside professional frameworks. Cyberdubbing shows that quality is no longer guaranteed by institutional authority alone, as audiences are capable of critical evaluation and willing to recognize non-professional alternatives when they better meet expectations of performance, style and cultural resonance. At the same time, this trust should be understood cautiously: active commenters represent only a minority of viewers and enthusiastic comments do not necessarily indicate long-term commitment or actual behavioral intention. Rather, they reveal a situated and relational form of trust shaped through comparison, participation and shared interpretation.

Declaration of AI use

In preparing this manuscript, the author used ChatGPT 5.5 for language editing. All AI-assisted content was reviewed, edited and verified by the author, who takes full responsibility for the accuracy and integrity of the final text.

References

- Baños, Rocío and Jorge Díaz-Cintas. 2024. "Exploring new forms of audiovisual translation in the age of digital media: Cybersubtitling and cyberdubbing". *The Translator* 30(1): 129–144.
- Cai, Jane. 2026. "Why are modern Chinese audiences shying away from dubbed films?". *South China Morning Post*, [May 2, 2026]. <https://www.scmp.com/news/china/politics/article/3352213/why-are-modern-chinese-film-audiences-shying-away-dubbed-films>
- Chaume, Frederic. 2012. *Audiovisual Translation: Dubbing*. London: Routledge.
- Chen, Xuemei. 2025. "Dubbing and reception: The revival of Chinese dubbese as internet memes on social media". *The Journal of Specialised Translation* 44: 85–104.
- Díaz Cintas, Jorge and Pilar Orero. 2010. "Voiceover and dubbing". In Yves Gambier and Luc van Doorslaer (eds.) *Handbook of Translation Studies*, pp. 441–445. Amsterdam and Philadelphia: John Benjamins.

- Du, Weijia. 2018. "Exchanging faces, matching voices: Dubbing foreign films in China". *Journal of Chinese Cinemas* 12(3): 285–299.
- Feng, Lin. 2025. "Voicing the West: Dubbing artists and invisible labor behind the socialist screen". In Lin Feng (ed.) *Film history and screen culture in and beyond Greater China*. Routledge. 72–95.
- Gottlieb, Henrik. 1998. "Subtitling". In M. Baker (ed.) *Routledge Encyclopedia of Translation Studies*, pp. 244–248. London: Routledge.
- Kress, Gunther and Theo van Leeuwen. 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold.
- López Salvador, Alicia and Giselle Spiteri Miggiani. 2025. "Creative freedom versus conventional norms in SDH: The case of Stranger Things". *Translation & Interpreting* 17 (2): 138–154.
- Massidda, Serenella. 2025. "Audiovisual translation and audience participation in the digital age". In S. Lu, S. Lu and L. Liang (eds) *Danmu-Mediated Communication and Audiovisual Translation in the Digital Age*. Routledge. 10–29.
- Massidda, Serenella. 2025. "Exploring creative audiovisual translation in the age of AI". *Translation Spaces*. <https://doi.org/10.1075/ts.25013.mas>
- Masunaga, Samantha. 2025. "'Zootopia' was a major hit in China. Will its sequel do as well?". *Los Angeles Times*. November 26. <https://www.latimes.com/entertainment-arts/business/story/2025-11-26/zootopia-was-major-hit-in-china-will-its-sequel-do-as-well>. Visited April 2026.
- Pym, Anthony. 2000. "Translating linguistic variation: parody and the creation of authenticity". *Traducción, metrópoli y diáspora*: 69–75.
- Pym, Anthony. 2015. "Translating as risk management". *Journal of Pragmatics* 85: 67–80.
- Pym, Anthony. 2020. "Trust-based translation history: Guideline questions and an illustration". *Chronotopos – A Journal of Translation History* 2(1–2): 146–160.
- Pym, Anthony. 2021. "Cooperation, risk, trust: A restatement of translator ethics". *STRIDON: Journal of Studies in Translation and Interpreting* 1(2): 5–24.
- Romero-Muñoz, Alejandro. 2023. "Multimodal analysis as a way to operationalise objectivity in audio description: A corpus-based study of Spanish series on Netflix". *Journal of Audiovisual Translation* 6 (2): 8–32. <https://doi.org/10.47476/jat.v6i2.2023.251>
- Zhang, Chunbai. 2004. "The translating of screenplays in mainland China". *Meta* 49(1): 182–192.
- Zhu, Ying. 2025. "Debate flares as Zootopia 2 unveils star-studded Chinese dub". *City News Service*. November 20. <https://www.citynewsservice.cn/articles/shanghaidaily/news/debate-flares-as-zootopia-2-unveils-star-studded-chinese-dub-5nq53q7n>. Visited April 2026.

“Art, esthetics, ideal”: an ethics for literary translators in the age of translating machines

GABRIEL GONZÁLEZ NÚÑEZ

The University of Texas Rio Grande Valley, United States

The advent of effective machine translation has raised several issues for translators, including the question of how a human translator ought to translate in this age of translating machines. In this context of new ethical questions, the present essay argues that an answer may be found in Pedro Figari’s (1912) philosophical essay on esthetics. Following Figari, the current essay argues that literary translation is art, that as such it can produce an esthetic emotion on readers and that the purpose of such art and esthetics is an ideal in the form of the continuous betterment of the human condition.

Keywords: machine translation, human translator, ethics, Figari, art, esthetics, ideal

“It seems unreasonable
to ask translators to save the world”
(Pym 2021: 21)

Introduction

The story is told of Diogenes who, after he was “forced to flee” or perhaps “banished” from the ancient Greek city of Sinope due to some irregularities with coinage, decided to live “a simple mode of life” (Laërtius 1853:224). The simple life made him arguably a difficult person and he often went about taunting others. Once he lit a candle, or perhaps a torch or a lamp, and went around in broad daylight saying he was “looking for a man” (Laërtius 1853:224). When Diogenes reportedly said “anthropon zeto,” he was intimating that the people around him lived in such a way that they could not be considered humans. His quest was in some ways ahead of its time, because centuries later, technology has advanced to the point that we might rightly ask what makes our own activity uniquely human (Medlir and Lamont 2021).

Classical philosophers such as Diogenes were concerned with an important question: “How should one live?” This is an ethical question about human activity. Along similar lines, ever since the emergence of Translation Studies as a field of scholarly inquiry, translation scholars have been interested in the ethical dimension of translational activity (Koskinen and Pokorn 2020). Such scholars have in essence been asking: “How should one translate?” (Lambert 2023:33). This question has been explored by theorists through different approaches, including those that value representation, service, communication and established norms (Chesterman 2001:139–142). While such ethical explorations have traditionally focused on translation itself, increasingly there is a focus on translator ethics, i.e., on how translators should behave as translators (e.g., Pym 2012). Ethical considerations on this front become more complex as machine learning technologies have yielded machines that can translate. The advent of these translation technologies has raised a number of issues surrounding ownership, confidentiality, collaboration, productivity and professional identity (Bowker 2020). It has also raised concerns about how translating machines—whether neural machine translation or large language models—ought to be designed so that the technology is human-centered, i.e., human-augmenting, and not machine-centered, i.e., human-replacing (Carpuat et al. 2025). In this brave new world of translating machines, many questions arise. The question of how a human translator ought to translate is one that merits exploration (Pym 2020:157).

The answer may well merit an entire book, but this short essay will address a narrower form of that question: “How should a human translate a literary text?” The argument might be made, of course, that the old ethical approaches still hold. While this may be true to some extent, in a context where machine translation is increasingly the norm, new questions beg for new answers. Some of them have to do with the ethical problems machine translation creates for literary translators. These are questions about ownership, confidentiality, collaboration, etc. There are, however, other types of questions. They have to do with the human dimension of translation in a machine translation world. What is it that a human brings to translation that a machine does not? Is there an ethical element that humans, because they are human, bring to the task of the translator? This is a set of questions about what it means to translate *as a human* now that translating machines have become commonplace. In short, what is the ethical way a human ought to translate because he or she is a human (not a machine)?

The human element of translation seems a lot more relevant in some areas of translation than in others. The highly formulaic instructions for assembling a dining room table at home? It probably doesn't matter whether the translating was done by an actual human or a machine imitating a human. A novel where a number of literary devices, including allegory and intertextuality, are deployed to explore the consequences of self-indulgence? It probably matters that a human translate this kind of texts, particularly

because the language through which these texts explore the human condition can produce potentially transformative esthetic effects. Faced with such texts, neural machine translation tends to produce flatter, more literal translations (Guerberof-Arenas and Toral 2022) and large language models can fail to properly handle cultural terms (Al Rousan et al. 2025) and can struggle when asked to work with low-resource languages (Peng et al. 2023). As Ayvazyan et al. have observed, “as soon as you get into long sentences, archaic diction and playful implicatures, the difficulty increases and the [translation] technologies falter” (2024:234). Thus, in areas of translation where the human element seems particularly relevant, such as when translating literature carefully crafted by humans in order to poignantly explore the human condition, it behooves us to ask what a human translator ought to do. Or at least, how he or she ought to go about the business of translating literature in a human way.

This essay will focus on addressing that last question. To do so, it will first address what has already been proposed in Translation Studies about ethical behavior, with a focus on the ethics of literary translation. Then the essay will argue that as pertains literary translation, a key to human translation may be found in Pedro Figari’s biological philosophy regarding esthetics. The essay will consequently argue that literary translation is art, that as such it can produce an esthetic emotion on readers and that the ethical imperative of such art and esthetics is an ideal in the form of the continuous betterment of the human condition.

Ethics: A (few) Translation Studies perspective(s)

Before proposing an ethics for literary translators in the age of translating machines, one must consider how scholars in Translation Studies have approached ethics. The first step in such a consideration should be agreeing on a basic understanding of the concept of ethics. On this point, José Lambert reminds us that ethics, at least as a branch of Western philosophy, has to do with “studying society as it should be rather than as it is, from the Greek ‘ethos’ meaning *custom, disposition*” (2023:14). Of course, when philosophers and scholars speak of ethics, they “mean not just any custom, habit, or character, but a desirable one” because it is *good* (Lambert 2023:14). In turn, Kaisa Koskinen and Nike Pokorn highlight that the term ethics is polysemic and it may refer to either a “code of moral rules,” a “system or theory of moral values or principles,” or “the philosophical study of morality” (2020:2). In these understandings, the common factor is that ethics is about the rules, system, or theory of morality. Incidentally, a distinction is sometimes observed between ethics as having to do with the collective and morals as having to do with the individual (Lambert 2023:13), but this is not always a clear distinction, as can be seen in the use above of morals to

describe ethics. Whatever the case may be, in the field of Translation Studies, ethics tends to be understood as “the subfield that aims to understand what is good and bad, right and wrong in *translatorial praxis*” (Koskinen and Pokorn 2020:3). Consequently, in terms of this essay, one may think of an ethical literary translator as one who has good or desirable customs, habits, or character when it comes to the practice of translating literature.

In the field of Translation Studies, a highly influential understanding of ethics comes from Andrew Chesterman (2001, 2020), who surveyed the main ethical currents within the field and proposed that they can all be understood as comprising an ethics based on virtue. A virtue-based ethics holds that one ought to seek to do good, which requires that one must first decide to be good (Chesterman 2020:13). What is good is determined by a number of values, which Chesterman outlines within Translation Studies to be truth (leading to an ethics of representation), loyalty (leading to an ethics of service), understanding (leading to an ethics of communication) and trust (leading to an ethics based on norms) (2001:139–142; see also Chesterman 2020:15–16). He has also identified some emerging values in the field, including justice and ecological responsibility (2020:20–22).

When considering virtue ethics, it is hard to argue that it is inadequate for a world in which humans translate alongside machines. Having values that are considered virtuous is, after all, a distinctly human trait that stems from our own self-awareness and consequent desire *to be*. Algorithms do not have values—only programming. Programming will predetermine functioning within very specific parameters, e.g., collating the work of human painters to create a new painting based on specific prompts. It is true that programming can tell the machine to behave in a way that its programmers might consider desirable. However, it cannot make the machine decide to be good by doing good because the machine has no consciousness, sentience, or soul (whichever of these overlapping terms one might prefer). Thus, ethics for literary translators in the age of translating machines must be a virtue-based ethics, one that requires the conscious decision to be good by doing good.

Because we are exploring humans as translators, it is helpful to move from general translation ethics to a focus specifically on the ethics of translators as such. Anthony Pym has defined “translator ethics” as “an ethics not of translations as things but of the people who make decisions concerning translations” (2020:147). This ethics is based on interpersonal relations, which limits it to the translator who engages with other humans as a translator (Pym 2020:147–148). It operates within “a framework of cross-cultural dialogue,” which is where Pym sees translators as exercising their agency (2012:166). In considering the work of translators as agents of such cross-cultural cooperation, Pym argues that the first ethical question is whether something should be translated at all (2012:12, 134–135). Overall, it is hard to contend against this starting point. The first ethical decision would logically be whether one should act, and in the case of a literary translator, whether they

should translate. Pym posits that the answer to this initial question is that the translator should engage in translation when it leads to cooperation, or mutual benefit, between cultures (Pym 2012:136-139). Once translators decide to translate, they should not do so at any cost. The benefits of cooperation should be weighed against the different costs of translation (Pym 2012:139-150). This entails some risk analysis, one that might take into account different strategies to reduce risk (Pym 2021:13-15).

Both Pym and Chesterman are concerned with how translators ought to make ethical choices as they go about their work. Chesterman bases his proposed ethics on philosophical notions about being good, while Pym bases his own proposed ethics on ideas found in neoclassical economics. For Chesterman, translation ethics ought to be about doing good in order to be good, while Pym argues that the ethics of translators ought to focus on cooperation because it leads to mutually beneficial outcomes. There is a great deal of overlap in these proposals. In fact, one might argue that Pym's cooperation ethics can fit well within Chesterman's virtue ethics, much the same way translator ethics can fit within the broader ethics of translation. Consider that Chesterman argues for an ethics based on values, but his list of values is open-ended. One might rightfully consider cooperation a value, and thus a translator that decides to engage in translation in order to advance cooperation between cultures is making a values-based decision: the translator decides to translate because they value cooperation. Where Chesterman and Pym's proposals seem to diverge is the motivation: Should one cooperate because it is a good thing good people do or cooperate because everyone will be better off through it? This distinction may make a difference, particularly on what cost the translator is willing to pay to facilitate that cooperation.

In addition to a general ethics of translation and a narrower ethics of translators, one might consider if ethics have been proposed for literary translators. In weighing this issue, Cecilia Alvstad concludes that "there is no comprehensive ethical framework which can help translators evaluate the consequences of different solutions to various ethical dilemmas" (2020:191). Even so, on surveying the Translation Studies literature that is relevant to literary translator ethics, she finds four main approaches.

One approach is based on Descriptive Translation Studies (DTS). According to Alvstad "although DTS provides no prescriptive advice on the ethical advantages and disadvantages of different solutions, translators will find abundant information in its descriptive studies on how other translators have solved similar problems" (2020:185). This would fall within Chesterman's value-based ethics. In essence, literary translators who value trust may look to DTS to know what norms they ought to follow in being trustworthy translators.

Another approach is the "ethics of difference" popularized by Lawrence Venuti and based on Antoine Berman and Friedrich Schleiermacher (Alvstad

2020:185-186). It too fits into the value-based ethics of Chesterman, who considers it an ethics of representation based on the value of truth.

An additional approach is found in George Steiner's hermeneutics, where translation is construed to be "a four-step hermeneutic motion": 1. the translator believes there is something that ought to be translated in the text, 2. the translator extracts meaning from the text, 3. the translator brings that meaning across to the target culture and 4. the translator directs the reader back to the source text (Alvstad 2020:182-183). It is in this fourth step where Steiner proposes an ethical imperative for literary translators: "The enactment of reciprocity in order to restore balance is the crux of the *métier* and morals of translation" (1998:300). This reciprocity is an ethical imperative in the sense that, inasmuch as the act of translating a literary text has extracted something from the text and added it to the target culture, that same act of translation must add something back to the source culture. Steiner (1998) proposes several strategies to create reciprocity or restore balance (e.g., by granting prestige to the source text) but ultimately finds that this ethical imperative is hard to fully realize and there often is imbalance between both sides. So it is that ethics is often an ideal toward which to aim. We will return to this idea.

Another ethical approach that literary translators may take is found in Henri Meschonnic's poetics (Alvstad 2020:181-182). Meschonnic defines ethics not "as a social responsibility, but as the pursuit of a subject striving to constitute itself through its activity, but where the activity of the subject is the activity by which another subject constitutes itself" (2011:35). Or, in plainer words: "Ethics is what one does with oneself, and with others" (Meschonnic 2011:45). Thus, an ethical act is one that transforms the subject, arguably for the better, and in the case of literature, that transformation is double: "the subject who writes and the one who reads" (Meschonnic 2011:36). In the case of translators, one might surmise the transformation also includes the one who translates. In Meschonnic's view, true translation only occurs when the target text does the same thing as the source text. This becomes an imperative for ethical translators: "we must do, with the means available in the target language, what the text has done to its source language" (Meschonnic 2011:85). This should not be confused with the functionalist approaches explained by Nord (2018), where purpose is the steering principle. Here, the text must act in the same way in both source and target languages, hence the importance of maintaining prosody, structure, wordplay, etc., as the totality of these things is what the text is doing. And how the text behaves is what allows it to effect transformations in people. At the heart of Meschonnic's insistence that an ethical literary translator will respect the poetics of the source text, i.e., how the text uses words to do what it does, is a crucial idea, namely, that texts can do things to people. They can transform individuals, and in the aggregate, societies. We will return to this idea as well.

Ethics for literary translators: A biological perspective

Grounded on the ethical considerations of previous translation scholars, this essay will now look sideways at esthetics in order to propose an ethics for literary translators. Specifically, it will suggest that Pedro Figari's (1912) philosophical essay on esthetics, *Arte, estética, ideal* [Art, Esthetics, Ideal] can provide ethical guidance in areas of translation where the human element seems particularly relevant, such as in the translation of literature. Before continuing, a word on Figari and his work seems necessary.

A turn-of-the-century Latin American intellectual, Pedro Figari Solari (Montevideo, 1861-1938) was a lawyer, politician, educator, author and painter (see Sanguinetti 2013). Today, he is mostly remembered for his work as an impressionist painter. Notably, Figari was *not* a translator or scholar of translation. His concerns were wide-ranging but focused greatly on art and esthetics.

In 1912, he published the first edition of his book *Arte, estética, ideal*, where he outlines some of the ideas that inform the present essay. The essay is considered a treatise on esthetics, but as Arturo Aldao points out, the scope is much broader (1960:vi). It is a work of "antropología filosófica" [philosophical anthropology], where the author's reflections are grounded on our nature as biological beings (Aldao 1960:vi-vii). Tellingly, the third French-language edition of the book changes the name to *Essai de philosophie biologique. Art, esthétique, idéal* [An Essay on Biological Philosophy: Art, Esthetics, Ideal] (1926). The definitive Spanish-language edition was published as a posthumous work that retains the original title and separates the lengthy essay into three volumes, one for each of the three elements in Figari's thought (1960). Even so, the key observation here is that Figari's philosophical proposal arises from his understanding of the human being as an "organismo biopsíquico" [biopsychic organism] that exists in the physical world (Aldao 1960:x). In other words, Figari's philosophy is rooted in the fact that as humans we are self-aware, biological beings that exist in physical surroundings. For Figari, that reality carries with it ethical duties, which he couches in the language of esthetics but that can be applied generally. This makes Figari's thinking pertinent to artistic activity, but the present essay will apply Figari's biological philosophy specifically to literary translation as carried out by humans.

To understand why Figari's thought can be applied to literary translation, one must first understand his expansive view of *arte*, or art. To him, art is not limited to the creation of esthetic products; rather, art is the application of biological intelligence to the improvement of our state in the world. This is not as unusual a definition as it may initially seem. At the dawn of the 20th century, one definition of the word *arte* was "todo lo que se hace por industria y habilidad del hombre" [all that is made through man's work and skill] (Real Academia Española 1899:96). Figari builds on this baseline definition. He

first explains that not every action undertaken by a human amounts to art. In his view the essential component of art is the deliberate application of intelligence toward improving our natural state (Figari 1960a:13-14). He illustrates this with a few examples: putting on a coat is not art, but weaving a sweater is; eating is not art, but making an arrow to hunt for food is; walking is not art, but engaging in a dance routine is (Figari 1960a:13). Notice that art is the *application* of intelligence toward a purpose, which can be distinguished from the actual product that may result from said application, e.g., the weaving of a sweater (art) as opposed to the sweater itself (the product of art). This definition of art moves well beyond the products of fine art, especially because it rejects the idea that art, to be art, cannot serve a practical purpose (Figari 1960a:35-39). In fact, Figari's conception of art encompasses science too, inasmuch as science is the systematic application of human intelligence to furthering knowledge for the benefit of humans (Figari 1960a:28-35). It becomes evident, therefore, that Figari's concept of art is large enough to encompass literary translation. Inasmuch as a literary translator deliberately applies their intelligence to solve the problem of communication, they are engaging in art and the end result, their translation, is a product of art.

What this concept of art does not include is the application of *machine* intelligence. Figari argues that art is a possibility for modern humans, primitive humans, and even higher species of animals, but he does not afford the same possibility to their tools. At this point it may be worth remembering that computers are indeed tools. In the end, what we metaphorically call "intelligence" in a machine is but a series of computations or mathematical operations. These are extremely complex operations that require immense computing power, but ultimately, they are the calculations of a tool. A highly sophisticated tool, but a tool nonetheless. Of course, in Figari's day the idea of a thinking tool was merely an exercise in fantastical imagination. Even so, it is doubtful Figari's biological philosophy could be applied to machines: our own self-awareness and the subsequent conscious decision to improve our (inter)actions are essential components to his understanding of art. The difference between an abacus, a calculator and a computer is one of degree, but even at their different degrees, one thing these machines have in common is that they are not self-aware, no matter how good some of them are at mimicking human behaviors.

In Figari's biological philosophy, art is a necessary stepping stone for humans to reach esthetics (Figari 1960b:179-180). As humans we apply our intelligence to create many things, but not everything we create is beautiful. In other words, not all art is esthetic. This is so because not all products of art can produce an esthetic reaction in us. For example, an arrowhead is a product of art that would not normally produce an esthetic reaction (Figari 1960b:178). Similarly, lines of programming code written by a human are a product of art that is not necessarily esthetic. Interestingly, not all esthetic emotion is the result of art either. For example, twilight in the countryside can

be perceived as esthetic, but it is not artistic because no human intelligence created it (Figari 1960b:178). Similarly, a machine can write a poem that imitates humans well enough to sound esthetic (Nair 2025), but because it is not the result of human intelligence, it is not artistic. In Figari's thought, what matters is not that esthetics exists in the world but that as humans we make things that can create esthetic reactions in others.

The concept of esthetic reaction is important to Figari's philosophy. He argues that as we go about, we employ a combination of both objective and subjective strategies to interact with our surroundings (Figari 1960b:49). Because of how we have evolved, we are predisposed to idealization (Figari 1960b:91) and that results in certain things causing in us "estados psíquicos que llegan a veces hasta la propia emoción intensa" [psychic states that can occasionally even become intense emotion] (Figari 1960b:87). We experience such emotional states, including esthetic emotion, because as human beings we are unable to process the world through reason alone (Figari 1960b:113). In essence, we cannot help but think *and* feel. From time to time, what we feel is esthetic. The nature of this esthetic emotion is such that as humans we intuitively seek it in a way similar to how we seek physical nourishment (Figari 1960b:92). Figari mentions as examples of esthetic emotion the feeling a person might experience observing a sculpted marble, watching a painted cloth, or hearing a series of structured sounds (1960b:87). In terms of translation, one might add that reading a combination of words on a page or a screen can produce an esthetic emotion. Therefore, a literary translation can be esthetic. This esthetic emotion, which is stirred within humans when experiencing certain things, is called beauty.

Where Figari's thought reaches an ethical dimension is in his concern for the ideal. Philosophically, the ideal—a major concern for many Uruguayan intellectuals in the early 20th century (Aldao 1960:xix)—can be understood as an ethical aspiration to higher forms. As espoused by fellow Uruguayan philosopher José Enrique Rodó, the ideal can be seen as "a striving towards the highest standards in human endeavor and values: good behavior towards the self and others; beauty in art; truth in intellectual work" (San Román 2018: 146). Figari sees the search for such an ideal as a consequence of our biological nature. As we interact with our surroundings, we are driven to gradually improve those interactions and consequently our place in the natural world (Figari 1960c:9-11). This leads to the ideal:

¿Qué es el ideal, pues? Es la aspiración a mejorar, determinada por el instinto orgánico en su empeño de adaptarse al ambiente natural. En ese esfuerzo de adaptación que se manifiesta de tan distintas maneras, el propósito es uniformemente el mismo *mejorar*. (Figari 1960c:15)

What, then, is the ideal? It is the aspiration to improve, as determined by an organic instinct seeking to adapt to its natural environment. This

effort to adapt is manifested in many different ways, but in all of them, the purpose is the same: *to improve*.

The ideal is ever elusive, as there is no final, fixed point where no further evolution toward improvement is possible (Figari 1960c:19-20). Just as importantly for purposes of this essay, there are as many ideals as there are human beings. For each, the ideal will depend upon their specific circumstances and the aspirations that might arise in them (Figari 1960c:19-20). However, those millions of aspirations will include, in their fulfillment, some kind of esthetic emotion (Figari 1960c:20). This is where the three ideas converge. Figari sees in that pursuit of *being better* a moral or ethical imperative: “¿qué puede ser más lógico, más moral ni superior que el aplicar su intelecto a llenar esa aspiración orgánica?” [what can be more logical, more moral or higher than the application of one’s intellect to fulfilling this organic aspiration?] (Figari 1960c:18-19). Thus, Figari’s proposed ethics are in essence an imperative to seek for the ideal.

Figari thought of his biological philosophy as applicable to any human endeavor capable of producing esthetic emotion. His ethics are generally applicable to artistic activity, which includes literary translation. (In other words, literary translation is a subset of artistic activity, the subset this chapter seeks to explore from an ethical standpoint.) Figari’s approach, as applied to literary translation, provides a moral imperative that is grounded on human biology. As human organisms, we seek to improve our interactions with the world. Literary translators do this when they translate (an activity that requires the application of intelligence) as a way to increase communication and ease the flow of ideas between peoples and cultures. Such literary translators are especially well positioned to create an esthetic emotion in their readers, which among other things, can result in a more impactful experience. When readers find a translation beautiful, they experience a pleasing emotion that is more memorable than no emotion at all. This experience need not be life changing—a kernel of impact will do, but the impact matters. The point here is that a literary translation can become transformative for a reader, not just because of the ideas it presents but also because of the emotion that the words can stir. With that understanding, the ideal, the ethical imperative, for the literary translator is to continually improve their skill so that they can better produce those esthetic emotions in others. Thus, the literary translator need not go out to save the world, but they are ethically bound to attempt to improve it.

Of course, this is a guiding principle more so than a set of instructions. Very specific questions can be asked regarding the application of this principle and the answer may admittedly vary depending on the specific circumstances. One might rightly wonder whose world the translator should seek to improve. In a general sense, the betterment sought is that of the world we all share: the human translator ought to seek to improve everyone’s world.

In a practical sense, not everyone is involved in a translational exchange, so the betterment can be limited at the very least to the individuals involved in such an exchange. Here, there is some alignment with Nord's (2018) call for loyalty and Pym's (2021) for cooperation: Figari's (1960) call for betterment of necessity requires that other humans be considered. Once the needs and the well-being of those other humans are taken into account, then it falls upon the literary translator to rely on their own experience and knowledge to determine how to craft a translation that can improve the human condition. It is precisely the human translator's experience as an organism in the physical world that can help them determine how to best deploy the transformative potential of language to create the necessary esthetic effects. And also, as will be seen below, to determine whether a translation should be carried out at all.

Applying the ethics: a few test studies

As has been stated above, Figari's biological philosophy is intended to apply to human activity generally and was not designed for literary translation alone. Therefore, it may be helpful to test whether this ethics can be fruitfully applied to literary translators. To that end, a few test cases will be analyzed in this section. These have been analyzed by other scholars through different lenses and offer the advantage of coming from the real world.

Case 1 will focus on translating *Mein Kampf*. The book is part autobiography, part political manifesto, which makes it literary in nature (uncomfortably so, in more ways than one). About the book, Pym says:

In some parts it is certainly expressive, manifesting a strong first-person character, as befits an autobiography. In other parts, it gives a vision of history and is thus referential. Finally, its overall function is undoubtedly to convert readers to the cause of National Socialism... (2023:58)

Under Figari's ethics, there is threshold question that the literary translator ought to ask: should I translate this at all? Here, consideration of the book's purpose should be had. The book, as stated above, is intended "undoubtedly to convert readers to the cause of National Socialism" and the translation could serve that same purpose. The literary translator then needs to weigh whether a translation with such a purpose would help improve the interactions of people with their environment, which no doubt includes other people. The answer to this question is no. The book is likely to be unhelpful—it has even proven to be hurtful—so the literary translator should pass on this task. The list of odious literary artifacts where one should decline to translate is long.

With other works, it may not be as simple as with, say, antisemitic literature. There are texts that may be perceived as helpful by some and

hurtful by others. Case 2 will focus on such a text by considering the translation of *En Cuba*. Part travel writing, part poetry, part interviews and part essay, the book is literary non-fiction. About the book, Nord says:

In his book *En Cuba*, written after a first visit to Cuba following the Revolution in 1959, the Nicaraguan priest Ernesto Cardenal presents a subjective, politically biased view of Cuban society. He is enthusiastic about the changes brought about by Fidel Castro's government. At no moment does he pretend to be objective, and the reader cannot fail to be impressed, even though they may not share Cardenal's views. (2008:116)

Once again, the threshold question is whether the translator should translate this at all. Different translators will come to different conclusions, and functionalist approaches grant that such will be the case sometimes. Every person is differently situated and may aspire in different ways to improve the world. One translator may find Cardenal's view of Cuba and Fidel Castro as unhelpful and likely damaging. Such a translator should not accept this task. Another translator may believe that the changes brought about by Castro in the island are positive and inspiring. Such a translator should accept this task.

After weighing whether the translation can result in betterment or improvement for the readers—and if the answer is yes—then the next question is how one should go about translating. In this sense, Figari's ethics do not propose that specific solutions be used (e.g., transposition) or that a single approach is preferable to others (e.g., foreignization). Rather, what is proposed is that the translator apply their intelligence toward a translation capable of producing an esthetic emotion in the reader. (It goes without saying that emotion does not mean sentimentality.) There is something poetic and engaging about *En Cuba*, and the literary translator should aim to produce a poetic and engaging text in their target language as well. How to create such a text will depend on varying contextual factors that ought to be considered individually and collectively, and that is where the human translator provides an expertise that machines cannot. This esthetic goal is independent of the stated purpose of the specific translation commission. In this ethics, the ultimate goal of an artistic activity, such as literary translation, is improving human interaction with the world.

A more recent example of an ethical dilemma is found in Case 3, which addresses the translation of "The Hill We Climb" into Dutch (but the scenario applies in any language other than English). About this poem, Lambert explains:

Amanda Gorman, an American poet and activist and (importantly here) a young, black, female spoken-word artist, became an international sensation after reading her poem 'The Hill We Climb' at US president Joe Biden's inauguration in January 2021. Shortly after the

inauguration, publishers scrambled to distribute the poem worldwide with a host of translations soon commissioned. (2023:169)

What followed was an international controversy when it was known that the translator commissioned to transfer the poem into Dutch was not a young, black, spoken-word poet, in other words, the translator did not share the original author's life experience. The controversy was such that the commissioned translator decided to step down from the project.

This case highlights the question of whether someone should take on a specific task. Under Figari's ethics, one ought first to consider whether the task of translating this poem will bring about an improvement in the translation's readers. Given the context of the poem, it is likely that most translators would conclude that translating this poem can potentially bring about some improvement for readers of the translation.

But what of the controversial question: should a translator pass on the translation if they do not share the author's experiential background? Here Figari's ethics require taking a step back and assessing whether the translator has the requisite skill to carry out that goal. If the translator does not, then the translator should not. The reason is simple: without such skill, the ethical goal that prompted taking on the translation is frustrated.

The point is that the translator should have the requisite translation competences to do this well. What is debatable is whether a similar life experience is necessary to develop the competences that will allow for the effective application of the translator's intelligence to the task at hand. In this particular case, however, one need not even get that far in the analysis. It is likely the original Dutch translator lacked at least some qualifications: they were not a poet (at least at that point), they had never published a translation and their English language skills were limited. A translator with such a profile is unlikely to perform the translation well if the aim is to be able to create the esthetic emotion that was so essential in this particular poem.

Figari's ethics of the ideal provide a way forward, however. A properly qualified translator should take on the task of pursuing the ideal of constant improvement of the world through translation. An unqualified translator, in turn, can choose to pursue the improvement of their own skill so that they can effectively carry out such tasks in the future. In the alternative, they may turn to other translation (or artistic) tasks. In other words, the ethics of the ideal calls on literary translators not only to seek to improve the interactions of others with the world but also their own. As the translator seeks to constantly improve their skill, their ability to perform literary translations will improve, which in turn increases the probability that their translations will create the desired esthetic emotions.

The three cases above help illustrate how Figari's ethics of the ideal can guide literary translators. The criticisms might be levelled that such an ethics provides only the broadest of directions. A threshold question is presented—

should one translate?—but if the answer is yes, then only a general notion of pursuing an ideal is presented. To be clear, that is what this ethics proposes. The ethical obligation to pursue the ideal in literary translation—through the application of human intelligence to deliberately improve our interactions with our surrounding world—is not intended to be a code of rules about when to apply specific solutions. Thus, the approach is not prescriptive in terms of what strategies to use but rather suggests that specific strategies are context dependent and that the betterment of the human situation ought to be the guiding principle behind those strategies, whatever they might be. Should the translation be more literal or free? Should the translation bring the reader to the text or the text to the reader? It depends on what will better help improve the reader's interactions with the world.

Conclusion

In sum, this essay is broadly concerned with human literary translators in a world of translating machines. The essay is built on the assumption that literary translation, when engaged in by humans, has an ethical dimension that we ought to take into account. It is on that ethical component where humans and machines diverge. Machines can be programmed to act in specific ways, but it is simply the application of rules expressed through mathematical computations. Because as humans we are self-aware, we think *and* feel in ways that machines do not. For that reason, human ethics cannot be reduced to mathematical computations. A human ethics of literary translation must take into account our nature *as* humans.

This is where Figari's "philosophical anthropology" can provide some illumination. In essence, he argues that our nature as living, biological beings makes us seek to improve our state through the betterment of our interactions with the world. That is to say that as humans we seek to improve our condition. When we apply our intelligence, through deliberate actions, to improving those interactions and, consequently, our own state, we are engaging in art (in a broad sense). The result of these efforts, in specific circumstances, can lead to an esthetic emotion, one which we both enjoy and seek after. The ideal is the constant, volitional effort to seek the improvement of our actions and circumstances so that the esthetic emotion may be realized. The pursuit of the ideal is an ethical obligation.

As regards literary translation, one may think of an ethical literary translator as one that has good or desirable customs, habits, or character when it comes to the practice of translating literature. For a literary translator, then, the desirable custom or character is applying their intelligence only in ways that improve our interactions with our surroundings and thus our human condition. This is an artistic endeavor which, in some cases, may result in an esthetic emotion. Ideally, this is what literary translators should seek after.

The ideal, and the ethical imperative, is for literary translators to continually seek to improve themselves and their art so that their translations can create esthetic emotions.

Such an ethics would be at odds with some of the ethical proposals that have been put forth in Translation Studies. It would not align well, for example, with those approaches that argue one should always translate in a specific way, for example, in the “ethics of difference” championed by scholars like Venuti (1995). An ethics of the ideal requires that every literary translation be approached with two questions in mind: ought this to be translated?, and if so, how can one translate this in a way that is most likely to produce an esthetic emotion that will improve, even if for a moment, the reader’s situation? The answer to that question may sometimes be to translate in order to highlight difference, but on other occasions it may be to translate in order to highlight similarity.

Because an ethics of the ideal does not result in an approach to literary translation that mandates always translating in the same way, it is partially incompatible with Meschonnic’s ethics of the poetic. Meschonnic finds that it is ethically imperative to respect the poetics of the source text because the totality of a text’s literary strategies and composition are what allow it to do what it does. However, this approach fails to realize that a set of strategies in one language may have a different effect in another language; in other words, respecting the poetics of the source text may result in a target text that fails to create the same esthetic emotion in the target readers. Here is where an ethics of the ideal can lead to different outcomes. A literary text, in order to produce in the target readers an esthetic emotion, may need to deviate from the poetics of the source text and adopt a different poetics. Even so, while both approaches may differ in their outcomes, they share an important philosophical element. They both agree that texts can produce changes in individuals and, by extension, in entire societies.

This ethics of the ideal also differs from Steiner’s “ethics of reciprocity” in that the restoration of balance does not seem to even be a consideration in Figari’s biological philosophy. He developed his philosophy with general human activity in mind, so it does not address issues such as balance and reciprocity. However, there is a place in Steiner’s reflections that aligns with Figari’s own thinking, namely, in the notion that there is an ideal toward which to aim. For Steiner, this is balance and for Figari it is the improvement of human action. Both ethics, in the end, point the literary translator toward a duty of ever striving for better.

An ethics of the ideal is not likely to fully align with utilitarian approaches to ethics, such as the “ethics of cooperation” proposed by Pym. But here there is ample room for overlap. In both ethical approaches, there is an initial question as to whether something ought to be translated at all. In a cooperation approach, if the answer is that the translation would lead to mutual benefits between cultures, then the literary translator should proceed.

In an approach of the ideal, the literary translator would look for a different answer. They would only translate if doing so would result in improving the interactions of humans with their surroundings, which includes other humans. Of course, if the result of the translation is increased cooperation between people, this qualifies as improved interactions, so the two ethical approaches will often lead to similar conclusions. However, an ethics of the ideal is broader in the kind of literary texts it will consider ethical to translate. Any text that can help improve the situation of a human in their specific context should be translated. This implies that the key calculation is not a cost-benefit analysis that posits the cost translation versus the benefit of cooperation but rather an assessment by the translator as to the extent to which he or she can actually bring about improvement. (The translator may decide that a modest, temporary improvement is enough justification.) Thus, what matters in the ethics of the ideal is that the translator strive toward continual improvement of both their own actions and those of the readers of their translations.

Where an ethics of the ideal may be most comfortably situated is in a context of virtue ethics as described by Chesterman. Ultimately, Chesterman's approach is the most wide-ranging, as it allows for different values and presumably different outcomes, in the search for the good. For some literary translators the guiding value might be truth, loyalty, understanding, trust, justice, or ecological responsibility. By bringing Figari's biological philosophy into ethical considerations, a new value can be proposed, that of the ideal. This is an ethics in which literary translators engage in a lifelong process of developing their art in the pursuit of esthetics for the betterment of the conditions of humans. It fits within a virtue-based ethics because it requires the human that is translating to decide to be good by doing good.

This matters in an age of translating machines. By pursuing "art, esthetics, ideal" as an ethics in literary translation, we focus our work as literary translators on that which makes us human: the continual search for the improvement of ourselves and other humans. This striving for the ideal in literary translation may very well stop Diogenes on his tracks: In the midst of all these translating machines, here are the humans; they are the translators constantly improving their craft to translate texts that will improve the ways in which they and others experience the world.

Declaration of AI use

The author did not use any AI-assisted tools in the preparation of this manuscript.

References

- Al Rousan, Rafat, Raghad Jaradat, and Mona Malkawi. 2025. “ChatGPT translation vs. human translation: An examination of a literary text”. *Cogent Social Sciences* 11(1).
<https://doi.org/10.1080/23311886.2025.2472916>
- Aldao, Arturo. 1960. “Prólogo”. In P. Figari, *Arte, estética, ideal. Tomo I*. Montevideo: Ministerio de Instrucción Pública y Previsión Social. v-xxxiv.
- Alvstad, Cecilia. 2020. “Literary translator ethics”. In K. Koskinen and N. K. Pokorn (eds.), *The Routledge Handbook of Translation and Ethics*. Abingdon: Routledge. 180–194.
- Ayvazyan, Nune, Yu Hao, and Anthony Pym. 2024. “Things to do in the translation class when technologies change: The case of generative AI”. In Y. Peng, H. Huang and D. Li (eds.), *New Advances in Translation Technology: New Frontiers in Translation Studies*. Singapore: Springer. 219-238.
- Carpuat, Marine, Omri Asscher, Kalika Bali, Luisa Bentivogli, Frédéric Blain, Lynne Bowker, Monojit Choudhury, Hal Daumé III, Kevin Duh, Ge Gao, Alvin Grissom II, Marzena Karpinska, Elaine C. Khoong, William D. Lewis, André F. T. Martins, Mary Nurminen, Douglas W. Oard, Maja Popovic, Michel Simard, and François Yvon. 2025. “An interdisciplinary approach to human-centered machine translation”. In C. Christodoulopoulos, T. Chakraborty, C. Rose, and V. Peng (eds.), *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*. Suzhou: Association for Computational Linguistics. 22859–22879.
- Chesterman, Andrew. 2001. “Proposal for a Hieronymic Oath”. *The Translator* 7(2):139–154.
- Chesterman, Andrew. 2020. “Virtue ethics in translation”. In K. Koskinen and N. K. Pokorn (eds.), *The Routledge Handbook of Translation and Ethics*. Abingdon: Routledge. 13–22.
- Figari, Pedro. 1912. *Arte, estética, ideal. Ensayo filosófico encarado de un nuevo punto de vista*. Montevideo: Imprenta Dornaleche.
- Figari, Pedro. 1926. *Essai de philosophie biologique. Art, esthétique, idéal*. Paris: Ed. Revue de L’Amérique Latine.
- Figari, Pedro. 1960a. *Arte, estética, ideal. Tomo I*. Montevideo: Ministerio de Instrucción Pública y Previsión Social.
- Figari, Pedro. 1960b. *Arte, estética, ideal. Tomo II*. Montevideo: Ministerio de Instrucción Pública y Previsión Social.
- Figari, Pedro. 1960c. *Arte, estética, ideal. Tomo III*. Montevideo: Ministerio de Instrucción Pública y Previsión Social.

- Guerberof-Arenas, Ana and Antonio Toral. 2022. "Creativity in translation: Machine translation as a constraint for literary texts". *Translation Spaces* 11(2): 184–212. <https://doi.org/10.1075/ts.21025.gue>
- Koskinen, Kaisa and Nike K. Pokorn. 2020. "Ethics and translation: An introduction". In K. Koskinen and N. K. Pokorn (eds.), *The Routledge Handbook of Translation and Ethics*. Abingdon: Routledge. 1–9.
- Laërtius, Diogenes. 1853. *The Lives and Opinions of Eminent Philosophers*. C. D. Yonge (trans.). London: Henry B. Bohn.
- Lambert, José. 2023. *Translation Ethics*. Abingdon: Routledge. <https://doi.org/10.4324/9781003148265>
- Mema, Medlir and Chris Lamont. 2021. "Pilot: Finding the human". *Age of AI*, April 20. Audio, 38:01. <https://www.ageofaipodcast.com/1727342/episodes/8367716-pilot-finding-the-human>
- Meschonnic, Henri. 2011. *Ethics and Politics of Translating*. P.-P. Boulanger (trans and ed). Amsterdam: John Benjamins. <https://doi.org/10.1075/btl.91>
- Nair, B. Monika, Pavithra S., Meera Karthikeyan Nair Sathikumari, Malathy R., S. Mahadevan, and R. Kanagaselvam. 2025. "Algorithmic esthetics: Machine learning for reconstructing poetic imagination in English literature". *2025 Tenth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*: Chennai. <https://doi.org/10.1109/ICONSTEM65670.2025.11374627>
- Nord, Christiane. 2018. *Translating as a Purposeful Activity: Functionalist Approaches Explained*. 2nd ed. Abingdon: Routledge. <https://doi.org/10.4324/9781351189354>
- Peng, Keqin, Liang Ding, Qihuang Zhong, Li Shen, Xuebo Liu, Min Zhang, Yuanxin Ouyang, and Dacheng Tao. 2023. "Towards making the most of ChatGPT for machine translation". In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 5622–5633, Singapore: Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.findings-emnlp.373>
- Pym, Anthony. 2012. *On Translator Ethics: Principles for Cross-Cultural Communication*. Amsterdam: John Benjamins. <https://doi.org/10.1075/btl.104>
- Pym, Anthony. 2020. "Translator ethics". In K. Koskinen and N. K. Pokorn (eds.), *The Routledge Handbook of Translation and Ethics*. Abingdon: Routledge. 147–161.
- Pym, Anthony. 2021. "Cooperation, risk, trust: A restatement of translator ethics". *Stridon* 1(2): 5–24.
- Pym, Anthony. 2023. *Exploring Translation Theories*. 3rd ed. Abingdon: Routledge. <https://doi.org/10.4324/9781003383130>
- Real Academia Española. 1899. *Diccionario de la lengua castellana*. 13th ed. Madrid: Imprenta de los Sres. Hernadno y Compañía.

- San Román, Gustavo. 2018. *A Companion to José Enrique Rodó*. Martlesham: Tamesis Books.
- Sanguinetti, Julio María. 2013. *El doctor Figari*. Montevideo: Ediciones Santillana.
- Steiner, George. 1998. *After Babel: Aspects of Language and Translation*. 3rd ed. Oxford: Oxford University Press.
- Venuti, Lawrence. 1995. *The Translator's Invisibility: A History of Translation*. Abingdon: Routledge.
<https://doi.org/10.4324/9780203360064>

How to find Paula D.? In quest of translators' identities in early 20th century Croatia

SNJEŽANA VESELICA MAJHUT

Zagreb University, Croatia⁴

Anthony Pym suggests that translation history should focus on translators as people (1998: 160) rather than the abstract “translator.” My research focuses on various modalities of presenting translators' identities in the peritext of “Zabavna biblioteka” (the “Entertaining Library”), a series focused on translated literature in early 20th century Croatia (1913-1941). In addition, three translators are selected for closer analysis in order to reconstruct a picture of various and fluid professional profiles of a heterogenous group of people translating for the “Library.” In order to do this, the data from the peritext are complemented with available biographical data.

Keywords: peritext, presentation of translators, translators' identities, professional profiles

Introduction

Anthony Pym (1998, 2000, 2009) has consistently argued that research into translation history would be more productive if the focus shifted from studying texts to studying translators as individuals. Discussing methodological challenges in translation history research and possible directions to pursue, Pym titles one chapter “Translators” and explains: “I am consciously using the term ‘translators’ in the plural, so as not to paint myself into the far more abstract term ‘the translator’ in the singular” (1998: 160). According to Pym, historical research into translation should pay at least some attention to “translators as people with flesh and blood bodies” (ibid.), rather

⁴ : The presentation of this paper at the 2026 NRTIS Conference was supported by funding from the development component of the Faculty of Humanities and Social Sciences within the framework of the University of Zagreb Programme Agreement for the programme period 2025–2029.

than to “the translator” in the singular, as an abstract anonymous entity. Unlike “the translator”, which often refers either to “the discursive figure that has produced a translation” (ibid.) or to a professional designation, a member of the profession (ibid.), “translators” refers to people who were active participants in translation events, but also lived their lives and were likely engaged in other forms of “multidiscursive mediation” (Pym 2009: 33).

Pym’s suggestion to use translators rather than texts as a starting point should be seen within his broader and consistent effort to “humanise translation history” (2009). The focus on translators, in his view, does not simply involve accumulating extensive bibliographic data on the translations produced by a particular translator or collecting dry biographical information. If the task of translation history is “to make narrative sense of those same relations between cultures, we require more than just raw data about texts, dates, places, and names. We must also be able to portray active people in the picture [...]” (Pym 2009: 24).

However, the task of uncovering translators from the past as “real people” with “flesh and blood bodies” is not easy, and sometimes not entirely feasible. The availability of richer data that enables us to portray translators in past periods largely depends on the range of social and professional roles they held in addition to their involvement in translation. As Pym (2009) and Paloposki (2016) note, translators in the past were usually not “monoprosessionals” in the modern sense. Thus, I am not the only one whose previous research into the history of translation has shown that lives of the people whose main or exclusive activity was translation tend to remain obscured (cf., Pym 2009, Paloposki 2016). Commenting on the situation in the Hispanic field, Pym (2009: 32) points out:

First, you discover that it is rarely easy to find out about translators as people, with their own life stories and evolving ideologies. [...] the long lists tend to be limited to bibliographical titles, dates of birth, and the occasional geographical reference. Apart from that, the cases most fleshed out tend to be translators who found fame wearing a different hat, as authors, political figures, polemicists, and so on.

The main aim of this article is to examine the modes of presenting literary translators in the peritext of translated texts from one series of translated fiction into Croatian in the early 20th century and to correlate these with their professional paths, in which translation played either a major or minor role. Beginning with their presentation in the peritext, I will proceed to uncover details about the “real life” existence of three translators, which may help build a more nuanced picture of how translatorship was practiced before the shift towards viewing translators as “monoprosessionals.” Given space limitations, the focus is on a particular publishing series that primarily published translations of literary works into Croatian (1913-1941). While researching this series, I was particularly intrigued by the variety of modes in

which translators' identities were presented in the peritext and sought to unravel the motivations behind the strategies used to conceal translators' identities.

The first part of the article presents a brief overview of the broader social and cultural context in which the selected series was published. This is followed by an outline of the profile and aims of the series and its founder, who was at the same time publisher, editor and, occasionally, translator of the titles published in the series. The second part discusses the modes of presenting translators' identities, with particular attention to the use of pseudonyms and initials as strategies for concealing translators' identities. Finally, three translators' professional profiles are selected for more detailed analysis, to demonstrate that, in terms of their professional backgrounds, translators of the period constituted a highly heterogeneous group.

Zabavna biblioteka and Nikola Andrić

“Zabavna biblioteka” (The Entertaining Library) was launched by Dr Nikola Andrić (1867–1942) in 1913, in the wake of the First World War, and consistently released two books a month, mainly translated texts, until 1941, that is, until the outbreak of the Second World War. In 1940, nothing was published and the last four volumes—an anthology of Polish poetry (1939), an anthology of Russian poetry (1939), an anthology of Italian poetry (1939) and an anthology of French poetry (1941)—were not edited by Andrić. All of this suggests that Andrić was prepared to end his publishing and editing activities. Therefore, the fact that “Zabavna biblioteka” ceased to exist in 1941 is not directly related to the outbreak of the Second World War and the political changes in Croatia, namely the establishment of the puppet state Nezavisna Država Hrvatska (The Independent State of Croatia). Simply put, it is more closely linked to the health condition of its founder, publisher and editor, the central figure of the entire enterprise, who died the following year.

The emergence of several publishing series focused on translated fiction in the early 20th century in Croatia can be attributed to changes in the broader political and social context. At the turn of the century, Austria-Hungary increasingly lagged behind rapidly developing European countries such as France, the United Kingdom and Germany, which became the European “center.” This is particularly evident in regions under Hungarian rule, such as Croatia. For example, at the beginning of the 20th century, 50% of the population in the northern parts of Croatia were illiterate, while in the south, that is, Dalmatia, the proportion of illiterate people reached 70% (Župan 2016: 273). According to Croatian historians Gross and Szabo (1992), the beginnings of modern civil society in Croatia can be observed in the 1860s and 1870s, decades marked by industrialization, urbanization and a substantial intensification of art and cultural activities. However, the full development of modern civil society in Croatia should be placed at the turn of the 20th

century. One important feature of modern civil societies is the radical transformation of reading practices, which shifted from a privileged elite activity to a mass phenomenon.

It is, therefore, not surprising that this period saw growing interest in publishing translated literature. I will mention a few other publishing series of translated fiction to illustrate this point: “Svjetska biblioteka” (World Library) (1916–1925), launched by Koprivnica-based publisher, printshop and bookshop owner Vinko Vošicki; “Nakladna knjižara M. Kelovića”, which also launched “Svjetska biblioteka” in 1916; and “Merkurova zabavna knjižnica” (Merkur’s Entertaining Library), launched by the Publishing Institute “Merkur” (1917–1919). This shows that publishers could rely on a substantial readership interested in foreign fiction in Croatian translation, as reading fiction was no longer reserved for the elite, who, at least in northern Croatia, mostly read in German.

However, the dominance of books in German remained strong long after Austria-Hungary ceased to exist, which is well-illustrated in an anonymous article published in 1936 in “15 dana” [15 Days], a journal dedicated to literature and arts:

Our bookshop windows are full of German books...Should a foreigner come to our city, he will, by the absolute number of German books in the windows of our bookshops, reasonably conclude that we do not have any original or translated literature, which is really disgraceful... (qtd. in Stipčević 2008: 97-98).

Thus, the above-mentioned publishing series aimed to provide access to world literature in Croatian. The reasons for launching “Zabavna biblioteka” are explained by its founder, publisher and editor, Nikola Andrić, in the “Catalogue of Zabavna biblioteka” compiled in 1931 to mark the publication of the 500th book in the series. In the text entitled “Instead of a Preface”, Andrić (1931: 5-6) states:

In February 1932, it will be nineteen years since "Zabavna biblioteka," without interruption, has reached Croatian readers. [...] On the first and fifteenth of each month, readers received a new novel. In this way, *the needs of those circles who were eager to obtain foreign editions were somewhat alleviated.* [...] “Zabavna biblioteka” records three particular achievements to its credit. The first is that *it brought many of the most prominent names of world literature into our literature.* The second: *that it presented all works (except Scandinavian and Turkish) in translations from the original languages;* and the third – that attention was paid, as much as possible, to the purity of the language. (my italics)

In the post-World War I period, unfavorable economic conditions led to the impoverishment of those sections of society who were the intended buyers of books. This is best illustrated by an extract from a speech by well-known

Croatian writer Dragutin Domjanić, president of Matica hrvatska, the leading cultural institution. In his speech at the General Assembly of Matica hrvatska in 1923, Domjanić said that immediately after the war people bought books because

everything was saturated with blood and everyone longed for something higher, purer and more spiritual. Those who could hardly afford to buy books sought books in order to relax, to find a relief...*but those times are over. Those who need a book can now hardly afford one.* Salaries of the more educated classes are shamefully low, especially if they are in the civil service, and the rich no longer need to pretend to have an interest in reading books. (my italics) (qtd. in Mandić 2014: 39)

The other publishers with similar publishing policies I mentioned, such as “Merkurova zabavna knjižnica” and “Svjetska biblioteka” Vinka Vošickog, could not overcome financial difficulties and after a while ceased to publish their series of translated fiction. However, “Zabavna biblioteka” managed to survive for almost three decades, largely owing to the skillful management of its editor. Its manifesto, “the best stories from world literature”, sometimes had to be adapted to circumstances. Thus, in addition to publishing first translations of a wide range of canonical works (by Joseph Conrad, Rudyard Kipling, Jaroslav Hašek, Pearl Buck, Knut Hamsun, etc.), Andrić also published popular works (by Maurice Decobra, Elinor Glyn, Edgar Wallace, etc.) that could guarantee commercial success.

Analysis of the presentation of the translators in the peritext of “Zabavna biblioteka”

In this section, various modes of presenting translators' identities are analyzed. Before I move on to the examination of how translators' identities are presented in the peritext of “Zabavna biblioteka”, I should provide some methodological details. Over a period of 28 years “Zabavna biblioteka” published 603 books, of which some 20 books were not translated but written by Croatian authors. It is also important to note that 603 books do not correspond to the publication of 603 literary works. Some larger works were published in more than one volume. For example, the translation of *Jerusalem* by Selma Lagerlöf was published in 1917 as four books (74,75, 76 and 77) within two volumes.

Within my broader research into the practices of publishing translated fiction in Croatia in the first half of the 20th century, I researched the entire series using both secondary sources, such as the descriptive bibliography of “Zabavna biblioteka” (Mandić Hekman 2014) and examining all the copies of the books stored in the archival collection in the Zagreb National and

University Library (NSK). The data on the modes of presenting translators in the peritext of the series were collected examining all the available copies of books published in the collection in the NSK. In order to see how the translators were presented in the peritext it was necessary to examine the copies of the books because this type of data is updated in the online catalogue of the NSK. For example, in the example I discuss below, in the peritext the translation was attributed to “B.D.”, but in the catalogue of the NSK, on the basis of the information obtained in the meantime, the translation is attributed to Branka Dizdarević. As my broader research focused both on the peritext and epitext of this series, using clues in secondary sources (Mandić Hekman 2010, 2014) I collected the relevant epitextual material, such as a few interviews by Andrić. With regard to the Catalogue published to mark the release of the 500th book I was lucky to obtain it from a second-hand bookshop in Zagreb.

The key elements relevant to this research include the ways of presenting the author of a source text (ST), the title of a ST and the title of a target text (TT), the identity of the publisher and editor of a TT and the identity of the translator of a TT. The cover usually contains the title of a Croatian TT, which in many cases is substantially modified compared to the title of the corresponding ST, often for marketing purposes. For example, in an interview in “Savremenik,” Andrić describes how he changed the title of Pierre Loti’s novel *Les Désenchantées* to *U haremu* (In Harem) and attributes the commercial success of the book to its attractive title.

The title page is usually organized as follows: at the top, the label of the publishing series appears in bold capital letters, ZABAVNA BIBLIOTEKA. Below it, information about the editor is written in smaller capital letters, “UREDJUJE DR. NIKOLA ANDRIĆ” (EDITED BY DR. NIKOLA ANDRIĆ), with Andrić’s name and academic degree in bold letters. As editor, Andrić always included his academic degree, thus emphasizing his symbolic capital in the literary field. The title of a TT is written in bold letters in a larger typeset, while the subtitle, in smaller letters, is given a less prominent position. The use of long subtitles, which indicated a genre or provided an explanation of the theme or context of a literary work, was a widespread convention in the 19th century. In this regard, “Zabavna biblioteka” was not an exception, although its subtitles were more elaborate than those in other similar series, suggesting that Andrić was well aware of their marketing potential.

The title pages are usually followed by forewords, which are most often, though not always, written by the editor. A foreword’s title is always simply the name of the author of the ST (for example, Charles Dickens, Alexei Kuprin), accompanied by his or her photograph. In line with this, forewords – sometimes four to five pages long – provide biographical data on the authors and a literary evaluation of their *oeuvre*. Additionally, the editor often provides information on previous translations of the selected author’s works,

as well as general statements about the role of translation and evaluative judgments of previous translations. Therefore, forewords in “Zabavna biblioteka” are a rich source of data for historical research on translation. However, their analysis is beyond the scope of this paper.

Looking at the copies of the TTs published in “Zabavna Biblioteka,” a translation scholar, particularly one interested in translation history, cannot help but ask: “Where are the translators?”

Up to 1932, information on the translator’s identity was placed at the end of the TT, which usually concluded with the phrase “preveo XY” [translated by XY], written in small letters. According to my insights, this used to be the norm in this period. Thus, in terms of the position and size of the typeset presenting data on all the contributors to the production of a TT, the information on the translator’s identity is clearly marginalized. As mentioned, from 1932 on, in “Zabavna biblioteka” this practice changed and the name of the translator along with the information on the source language was placed on the left side of the title page, usually in the phrase “s engleskog preveo Iso Velikanović” (translated from English by Iso Velikanović).

Still, despite the marginalization of translators’ roles in the peritext, for the entire enterprise to run smoothly and to deliver two translated books a month, it was necessary to gather and manage a large number of translators.

Transparently presented translators

In most TTs, translators are presented with their full names, sometimes accompanied by their professional attributes and academic titles. It should also be noted that Andrić himself translated several books, in whose peritext this is most often indicated as follows: “preveo Nikola Andrić” (translated by Nikola Andrić). However, in the first book in the series, a translation of several stories by Dostoevsky published under the title *Krotko janje* (1913) (A Meek Lamb), Andrić is presented as a translator in the following way: “preveo dr. Andrić” (translated by Dr Andrić). The second book in the series, the translation of Knut Hamsun’s *Victoria: en kærligheds historie*, was also done by Andrić, this time presented simply as “Nikola Andrić,” without an academic title. The practice of presenting his authorship of translations without an academic title was used in the peritext of Andrić’s translations from 1917 onward.

As for other translators who signed their translations with their full names, in some cases academic titles were used. For example, at the end of the translation of *Norina* (1913), a novel by Hungarian author Renée Erdős, it is indicated that it was translated by “Dr Ivan Lulić.” Similar examples include the translation of Haggard’s novel *She* (1916), titled *Ona. Misteriozna povijest jedne pustolovine* (She. A Mysterious History of One Adventure), in whose peritext the translation is attributed to Dr Mato Lisičar and Vanda Ibler-Novosel. The translation of *Misunderstood* (1917) by Florence

Montgomery is attributed to Dr Milan Bogdanović, a well-known translator of Shakespeare's plays into Croatian directly from English.

This practice shows that translators were highly educated and had earned academic degrees, though in various disciplines, not necessarily in philology and foreign languages. For example, the above-mentioned Dr Mato Lisić obtained his degree in law in 1906, after which he worked as a solicitor and notary public. According to the *Croatian Biographical Lexicon*, he contributed to a large number of periodicals, publishing translations, literary criticism, reviews, poems, essays and articles on cultural and political topics. Therefore, it seems that a range of his various activities, in addition to translation, may be a reason for his name, accompanied with an academic degree, to appear clearly in the peritext.

Anonymous translators

During its initial years, "Zabavna biblioteka" published one book per month. As reader interest grew, Nikola Andrić increased production to two books monthly, resulting in a total of 603 books across the entire series. The vast majority (approximately 580 books) are translated works, among which only 15 were published anonymously. Table 1 presents the bibliographic data for the TTs whose peritexts omit translator attribution.

It is immediately apparent that anonymous translations were published in the early years of the series, suggesting that it was not easy to gather enough translators to ensure the continuity of the series. Ivana Mandić Hekman (2010: 12), based on her review of Andrić's correspondence with Vatroslav Jagić, an outstanding philologist and Andrić's supervisor during his Vienna days, assumes that most of the anonymous translations were done by Andrić himself. However, the selection of the source languages for these TTs points to another aspect. In the forewords he wrote for "Zabavna biblioteka" editions, Andrić repeatedly stated that he insisted on publishing direct translations, which was a gesture of resistance –partially grounded in political animosity– against the widespread previous practice of using German translations as mediating texts. However, his foreign language skills included German, French and Russian, which strongly suggests that the anonymous translations listed above were indirect translations, most likely based on German mediating texts. This fact had to be concealed, although, as seen in the quotation from his text in the Catalogue (1931: 6), he eventually admitted that Scandinavian source texts were translated indirectly. However, he probably did not want his name openly associated with such translations.

Table 1: A list of TTs with no translators' identities provided

ST author	ST title and the year of first publication	ST language	TT title and the year of publishing
Wilde, Oscar	<i>The Canterville Ghost</i>	English	<i>Sablast od Cantervilla: realistično-idealistična pripovijest</i> (1913)
Neera	<i>Il libro di mio figlio</i> (1891), <i>Senio</i> (1892)	Italian	<i>Senio i Knjiga moga sina</i> (1913)
Croker, Bithia Mary	<i>The company's servant: a romance of southern India</i> (1907)	English	<i>Indijska plesačica. Roman iz engleskog društva južne Indije</i> (1913)
Danilevsky, Grigory	<i>Knjažna Tarakanova</i> (1883)	Russian	<i>Kneginja Tarakanova. Historijski roman iz 18. vijeka</i> (1913)
Feuillet, Octave	<i>La morte</i> (1886)	French	<i>Jadna Alijeta Roman iz aristokratskog svijeta</i> (1913)
Af Geijerstam, Gustaf	<i>Kvinnokraft</i> (1901)	Swedish	<i>Što žene mogu. Roman iz skandinavskog života</i> (1914)
Bang, Herman	<i>Ellen Urne. Skuespil i fire Akter</i> (efter Fædra) (1885)	Danish	<i>Grofica Urna. Roman jedne nesretne žene</i> (1914)
Wells, George Herbert	<i>The Invisible Man</i> (1897)	English	<i>Nevidljivi čovjek: groteskni i fantastični roman</i> (1914)
Glyn, Elinor	<i>His Hour</i> (1910)	English	<i>Tamara. Roman iz ruskog aristokratskog svijeta</i> (1914)
Couperus, Luis	<i>Noodlot</i> (1890)	Dutch	<i>Sudbina. Roman nesretnih ljubavnika</i> (1915)
Glyn, Elinor	<i>Three Weeks</i> (1907)	English	<i>Tri nedjelje. Roman iz ruskog dvorskog svijeta</i> (1915)
Nagrodskaya, Yevdokiya	<i>Gnev Dionisa</i> (1910)	Russian	<i>Dionisijev gnjev. Roman jedne slikarice</i> (1916)
Hamsun, Knut	<i>Pan: Af Løjtnant Thomas Glahns Papirer</i> (1894)	Norwegian	<i>Pan Iz zapisaka poručnika Glahna</i> (1917)
Guglielminetti, Amalia	<i>La rivincita del maschio</i> (1923)	Italian	<i>Muškarčeva osveta. Roman iz talijanskog života</i> (1925)

Translators presented with pseudonyms, codes and initials

Of the 580 translated works, 22 were attributed via initials or pseudonyms. Table 2 presents the bibliographic data for the TTs that utilize these pseudonyms and initials within their peritexts.

Table 2: A list of TTs with translators' identities provided as pseudonyms and initials with their real identities added in brackets

ST author	ST title and the year of first publication	ST language	TT title and the year of publishing	Translators' initials, codes and pseudonyms
Tikhonov, Aleksey	Dobej ego!	Russian	<i>Police verso Paralele</i> (1917)	Nik. N. (Nikola Nikolajević)
Glyn, Elinor	<i>Beyond the Rocks</i>	English	<i>Teodora: roman jedne mlade gospodje</i> (1917)	Š.B. (Branko Širola)
Bourget, Paul	<i>André Cornelis</i>	French	<i>André Cornelis: kriminalni roman</i> (1920)	Paula D.
Farrère, Claude	<i>Les civilisés</i>	French	<i>Civilizirani ljudi: roman sa krajnjeg istoka Azije</i> (1921)	Zlatko Tujjin (Zlatko Gorjan)
De Maupasant, Guy; Mirbeau, Octave, Daudet, Alphonse	No title is provided.	French	<i>Francuske novele</i> (1921)	Rakošev Pekarić (Pavao Rakoš)
Carco, Francis	<i>Verotchka l'étrangère</i>	French	<i>Tudjinka Vjeročka: roman iz ruskoga društva u Parizu</i> (1924)	R.M. Ivanović (Rudolf Maixner)
Michaëlis, Karin	<i>Den farlige Alder</i> (1910)	Danish	<i>Žena u opasnim godinama: listovi i dnevni zapisi jedne rastavljene žene</i> (1924)	Miloje Fruškogorac (Nikola Andrić)
Farrère, Claude	<i>Les petites alliées</i> (1910)	French	<i>Male saveznice: roman iz života francuskih pomorskih časnika</i> (1924)	Z.D. (Zvonimir Doroghy)
Maurois, André	<i>Ariel ou la Vie de Shelley</i> (1923)	French	<i>Ariel ili Život Shelley-a</i> (1925)	R.M. Ivanović (Rudolf Maixner)
Mauriac, François	<i>Le fleuve de feu</i> (1923) and <i>Le baiser au lépreux</i> (1922)	French	<i>Plamena rijeka: roman jedne grešne djevojke</i> (1925)	R.M. Ivanović (Rudolf Maixner)
Wied, Gustave	<i>Faedrene aede druer</i> (1908)	Danish	<i>Oci jedoše kiselo groždje: roman jednog vlasteoskog gnijezda</i> (1925)	Milan Zorić (Janko Ibler)
Kraigher, Alojz	<i>Mlada ljubezen</i> (1917)	Slovenian	<i>Mlada ljubav: roman iz ljubljanskog života</i> (1925)	Z.D. (Zvonimir Doroghy)
Niedźwiecki, Zygmunt	<i>Erotyki opowiadania</i> (1904)	Polish	<i>Erotika: ljubavne novele</i> (1925)	J.V. (Vatroslav Jagić)
Stanley, Henry Morton	<i>How I Found Livingstone</i> (1872)	English	<i>Kako sam našao Livingstone: izvještaj o ekspediciji po središnjoj Africi: (ilustrirano izdanje)</i> (1925)	M.L. (Martin Lovrenčević)

Dekobra, Maurice	<i>Gondole aux chimères</i> (1926)	French	<i>Gondola krilatih nemani. Kozmopolitski roman.</i> (1926)	S.A. (not identified)
Kennedy, Margaret	<i>The Constant Nymph</i> (1924)	English	<i>Vjerna nimfa: roman jedne umjetničke porodice</i> (1928)	D.Š. (Ivo Šrepel)
Curwood, James Oliver	<i>Swift Lightning</i>	English	<i>Zečja omča - kosa djevojačka: roman iz arktičkih krajeva sjeverne Amerike</i> (1930)	D.Š. (Ivo Šrepel)
Wallace, Edgar	<i>The Strange Countess</i>	English	<i>Zagonetna grofica: roman iz londonskog života</i> (1931)	E.K. (not identified)
Curwood, James Oliver	<i>The Last Frontier</i>	English	<i>Na kraju svijeta: roman sa krajnjeg sjevera Amerike</i> (1933)	M.N. (not identified)
Grey, Zane	<i>The Mysterious Rider.</i>	English	<i>Tajanstveni jahač: ljubavni roman s američkog Divljeg zapada</i> (1934)	V.D.S. (Viktor Dragutin Sonnenfeld)
Durtain, Luc	<i>Frank et Marjorie</i>	French	<i>Frank i Marjorie: roman o propasti puritanske i materijalističke kulture u Americi</i> (1935)	B.D. (Branka Dizdarević)
London, Jack	<i>Hearts of Three</i>	English	<i>U tajanstvenom carstvu: fantazija o traženju blaga starih Maja</i> (1935)	V.D.S. (Viktor Draguin Sonnenfeld)

I must honestly admit that I was not able to determine who is behind the pseudonym Paula D., used in the title of this paper. M.N., E.K. and S.A. also remain unidentified. The main source I used is the fundamental bibliographic list of pseudonyms and codes (Vidačić 1951). Vidačić listed pseudonyms and codes and linked them to real identities when they appeared regularly in periodicals of the time. Therefore, these were probably one-time translators.

Profiles of translators

Due to space constraints, this section focuses on outlining the professional profiles of three translators who authored multiple TTs published in “Zabavna biblioteka.” While two of them operated under concealed identities, the third not only signed the translation with his full name but also detailed his personal and professional trajectory within the peritext. These profiles offer a fragmentary yet multi-layered portrait of translation practitioners in early 20th-century Croatia, demonstrating how their translational work intersected with their other roles as cultural mediators.

Ivo Šrepel (D.Š.)

Ivo Šrepel (1899–1945), who signed as D.Š., produced two translations for “Zabavna biblioteka” within a short period. Both were translated from English: *The Constant Nymph* (1928), then a highly popular novel by Margaret Kennedy, published as *Vjerna nimfa: roman jedne umjetničke porodice* (The Faithful Nymph: A Novel about One Artistic Family) and *Swift Lightning* (1930) by James Oliver Curwood, published as *Zečja omča – kosa djevojačka: roman iz arktičkih krajeva sjeverne Amerike* (Rabbit’s Noose – Maiden’s Hair: A Novel from the Arctic Parts of Northern America).

According to available biographical data, Šrepel studied philosophy in Zagreb, Vienna, Prague and Paris. After earning a degree in esthetics and art history from the Sorbonne in 1921, he continued to live in Paris, where he became deeply involved in the French theatrical scene, establishing personal contacts with several French authors and translating French theater texts for staging in Zagreb. Upon returning to Zagreb, he worked as an administrative officer at the Croatian National Theatre and as a presenter at the newly-established Croatian Radio. From 1929 to 1940, he regularly contributed to “Jutarnji list” and other dailies with reports on fine arts events. From 1940 to 1944, under the regime of the Independent State of Croatia, he served as director of the Modern Gallery. Accused of using his position at the Modern Gallery to send material aid to partisan units for their promotional activities, Šrepel was dismissed and imprisoned in 1944. The circumstances of his death in 1945 have remained unclear.

Šrepel began translating during his early days in Paris, publishing translations of Baudelaire’s prose poems in 1920 under his full name, which brought him recognition in literary circles in Croatia. During his studies in Paris, “he translated intensively, which enabled him to become friends with several French authors whose dramas were staged in Paris” (Jirsak 2010: 13). This demonstrates that, for Šrepel, translation was a discursive practice combined with other mediating activities. This is evident from a document in Šrepel’s personal legacy (Vončina 1988:109), in which Josip Kulundžić, a Croatian playwright, “authorizes Mr. Ivo Šrepel, a writer and translator, resident in Paris, Rue Rennequin no. 63, to act on his behalf in all matters related to the staging of *Ponoć* (Midnight) on one of the stages in Paris.” As compensation for Šrepel’s translation of the play into French, Kulundžić agreed to share his authorial fees with him. In summary, for Šrepel, translation was just one of several activities he undertook as an intercultural mediator – a person who, due to his physical mobility and social skills, participated in various stages of cross-cultural transfer, both from cultural centers to Croatia and from Croatia to cultural centers. Along this path, he met Nikola Andrić and his circle, producing the two translations for “Zabavna biblioteka,” which he did not sign with his full name, as they were not considered to be prestigious.

Viktor Dragutin Sonnenfeld (V.D.S.)

Translations of Zane Grey's *The Mysterious Rider* (Tajanstveni jahač: ljubavni roman s američkog Divljeg zapada, 1934) and Jack London's *Hearts of Three* (U tajanstvenom carstvu: fantazija o traženju blaga starih Maja, 1935) are attributed to V.D.S., the initials used by Viktor Dragutin Sonnenfeld.

The first translation Sonnenfeld published was Schopenhauer's treatise *Über die Weiber* in 1922 in "Svjetska biblioteka" Vinka Vošickog. In the following years, he published several translations of Schopenhauer's works, as well as translations of other philosophical texts. In an article in *Hrvatska enciklopedija*, Sonnenfeld is described as "a Croatian philosophy writer and translator" who "translated fundamental philosophical works from German and English (F. Bacon, J. Locke, I. Kant, J. G. Fichte, F. W. J. Schelling, G. W. F. Hegel, A. Schopenhauer), thus making a substantial contribution to the enrichment and standardization of Croatian philosophical terminology." All his translations of philosophical works are attributed to him under his full name. However, it appears that Sonnenfeld, in addition to translating philosophical works for which he received public recognition, developed an interest in translating popular literature, but never signed such translations with his full name. A year after the translation of Jack London's *Hearts of Three*, in 1936, Sonnenfeld, who lived in Osijek and worked as an editor at the right-wing daily "Hrvatski list" (Croatian Paper), launched "Biblioteka Hrvatskog lista," a series publishing his translations of popular literature, always signed V.D.S. The motivation behind his venture into translating and publishing popular literature remains unclear: whether it was material interests or a need for relaxation from translating philosophy.

Another relevant element in understanding how he became involved with "Zabavna biblioteka" is, as in the case of Šrepel, his connection to the theatrical scene, a social circle where Andrić was well-connected.

Joe S. Rendulić

Of the translators whose identities are not concealed but highlighted in the peritext, special attention should be given to Joe S. Rendulić (1890–1970). He is the only translator collaborating with "Zabavna biblioteka" who was given the opportunity to present his personality in the peritext. Moreover, his foreword to *Anglosaske pustolovine: zbirka istinitih doživljaja* (*Anglosaxon Adventures: A Collection of Real Events*), published in 1926, is accompanied by his photograph. In the series, there are only a few translators' prefaces, mainly by translators who were established literary critics (Rudolf Maixner, Zdenka Marković, Bogdan Radica) and, in those cases, they comment on the ST author and evaluate the published literary work.

Rendulić's preface is notable because he uses it to present his life and explicitly position himself as an intercultural mediator. This is evident from the very beginning of the preface, where Rendulić explains his motives for translating the texts that he not only translated but also selected:

As a Slav – a conscious Croat – and a Yugoslav, I have spent a large part of my life in America, Australia, Africa, and Asia and have come into contact with these [Anglo Saxon] people and have learned to have a deep respect for them. With their energy, diligence, honesty, and ability to sacrifice, they do not possess that egotistical and cold “prone to calculation,” which is often attributed to them in Europe. (Rendulić 1926: 5)

As seen in the preface, Rendulić decided “to compile this collection of stories” (ibid.), all based on real events, which will enable “our people to come to know this race” (ibid.).

Regarding the translation itself, Rendulić points out that the source texts are “written in jargons, of which less than twenty percent could be found in lexicons” (1926: 6). Because he lived in those countries, his “literary conscience” is at peace, as he did his best (ibid.). A particularly interesting part of the preface is an account of Rendulić's life, which I will summarize below. Having spent his youth as a sailor in Patagonia, at the outbreak of the First World War he was in Australia, where he ended up in internment. A list of occupations he had before temporarily returning to Croatia includes “miner, actor, boxer, teacher, fisherman, insurance agent,” until he “became what I am right now: an author looking for a publisher” (Rendulić 1926: 8).

It is worth noting that Rendulić translated another six novels for “Zabavna biblioteka” between 1926 and 1931 and, after that, no further evidence of his literary or translation activity in Croatia could be found. According to a short article in *Hrvatska enciklopedija*, he died in Uruguay. Searching for additional data on him, I found on a website “Zorrilla Subastas”, an auction house in Montevideo, pictures of several landscape oils on canvas entitled “Croatian port” signed by his name.

Conclusion

This article presents a contribution to the efforts to humanize translation history by focusing on translators. The first step is an analysis of peritextual attribution of translations published in “Zabavna biblioteka”, a long-term publishing project that continuously published translated fiction into Croatian from 1913 to 1941. The usual practice of attributing translations in the series was to sign the translator's name in small font on the last page of the TT, which shows that, unlike other agents involved in the project, they occupied a marginal position in the peritext. While the majority of translations were

attributed to translators signed with their full names and even academic titles, 15 translations were anonymous and 22 were signed by translators' pseudonyms or initials, showing that strategies of concealing the real identities of translators were used. As for motivations lying behind concealing real identities of the translators involved, a number of reasons may be proposed. In the case of Nikola Andrić, the publisher and editor of "Zabavna biblioteka", who occasionally produced translations and often signed them with his full name, the motivation for his decision to sometimes remain anonymous may lie in the fact that his use of mediating texts was in strong contradiction with the norm of direct translation that he often proclaimed in his editorial forewords to the books published in "Zabavna biblioteka" and the collected epitextual material.

Within a heterogeneous group of people who were engaged as translators in "Zabavna biblioteka" three translators were selected for closer analysis. Biographical data on Ivo Šrepel, a prominent figure in the cultural life of the time, show that for Šrepel translation was a form of multidiscursive mediation he, owing to his physical mobility and social skills, combined with other cross-cultural mediating activities. On the other hand, the reason for using initials in the peritext of the particular translations from English in "Zabavna biblioteka" may lie in the choice of the TTs, which were not as prestigious as he would expect, since his early translation of Baudelaire was signed with his full name. Another interesting case is Viktor Dragutin Sonnenfeld, who gained public recognition as a translator of major philosophical works into Croatian, always signed by his full name. However, at the same time he translated popular fiction using initials instead of his full name. Therefore, in his case, the reason for not using the full name may lie in low prestige of such translation activity, which would not contribute to his status of "serious" translator and the symbolic capital it entailed. Unlike Andrić, Šrepel and Sonnenfeld, who chose to conceal their identities, Joe S. Rendulić signed all the translations for "Zabavna biblioteka" with his full name. In stark contrast to all the other translators engaged in the series, Rendulić's personal life is presented in the peritext of the first translation produced for the series. Prior to being engaged as a translator Rendulić changed a whole range of occupations, and for a short period stayed in Croatia, where he probably initiated the first translations of Zane Grey's works into Croatian, published in "Zabavna biblioteka."

To sum up, the patterns of concealing or revealing translators' real identities in this particular series point to a number of factors at play: for those who, along with translation, were engaged in a number of discursive mediating practices bringing public recognition (editing, journalism, publishing, intercultural mediation), the motives for concealing their real identities could be linked to not wishing to be publicly associated with translating low prestige texts (Šrepel and Sonnenfeld) or using indirect

translation, which is in contradiction with his publicly advocating the practice of direct translation as a desirable norm (Andrić).

To conclude details of the “real life” existence of three translators, selected for closer analysis, reveal, on the one hand, a more nuanced picture of how translatorship was practiced before the shift towards viewing translators as “monopersonals.” Further, focusing on translators as people reveals how they managed professional anxieties, such as the potential damage that indirect translation or low-prestige activities—like translating popular fiction—could inflict on their public image.

Declaration of AI use

The author did not use any AI-assisted tools in the preparation of this manuscript.

References

- Andrić, Nikola. (ed). 1931. *Katalog Zabavne biblioteke*. Zagreb: Naklada Zaklade tiskare Narodnih novina.
- Andrić, Nikola. 1936. “Nikola Andrić o sebi (Razgovor s Jubilarcem)”. *Savremenik* 26(2): 444-447.
- Gross, Mirjana and Szabo, Agneza. 1992. *Prema modernom građanskom društvu: Društveni razvoj u civilnoj Hrvatskoj i Slavoniji šezdesetih i sedamdesetih godina 19. stoljeća*. Zagreb: Globus.
- Jirsak, Libuše. 2010. *Ivo Šrepić: Dokumenti, vrijeme, sudbina*. Zagreb: Moderna galerija.
- Lisićar, Matija. *Hrvatski biografski leksikon (1983–2025), mrežno izdanje*. Leksikografski zavod Miroslav Krleža, 2026. <https://hbl.lzmk.hr/clanak/lisicar-matija>. Visited April 2026.
- Mandić Hekman, Ivana. 2010. *Leksikon Zabavne biblioteke: pisci, djela, književnosti: 1913.-1941.* Zagreb: Ex libris.
- Mandić Hekman, Ivana. 2014. *Knjiga o knjigama: Zabavna biblioteka 1913-1941*. Zagreb: Ex libris.
- Paloposki, Outi. 2016. “Translating and translators before the professional project”. *The Journal of Specialised Translation*, 25: 15-31.
- Pym, Anthony. 1998. *Method in Translation History*. Manchester: St. Jerome Publishing.
- Pym, Anthony. 2000. *Negotiating the Frontier: Translators and Intercultures in Hispanic History*. Manchester St. Jerome Publishing.
- Pym, Anthony. 2009. “Humanizing Translation History”. *Hermes – Journal of Language and Communication Studies*, 42: 23-48.

- Sonnenfeld, Viktor Dragutin. *Hrvatska enciklopedija*, mrežno izdanje. Leksikografski zavod Miroslav Krleža, 2013. – 2026. <https://www.enciklopedija.hr/clanak/sonnenfeld-viktor-dragutin>. Accessed April 2026.
- Vidačić, Marcel. 1951. "Pseudonimi, šifre i znakovi pisaca iz hrvatske književnosti". U: Barac, Antun (ed.) *Građa za povijest književnosti hrvatske*, 21: 7-141. Zagreb: JAZU
- Vončina, Nikola. 1988. *Hrvatska radio-drama do 1957.: komparativni pregled*. Zagreb: Školska knjiga
- Župan, Dinko. 2016. "Kulturni i intelektualni razvoj u Hrvatskoj u 'dugom' 19. Stoljeću". In Švoger, Vlasta and Turkalj, Jasna (eds.). *Temelji moderne Hrvatske. Hrvatske zemlje u "dugom" 19. stoljeću*. Zagreb: Matica hrvatska. 273-308
- Stipčević, Aleksandar (2008). *Socijalna povijest knjige u Hrvata* (book III). Zagreb: Školska knjiga.

On interpreting v. translation

KAYOKO TAKEDA

Rikkyo University, Japan

Conversations with Anthony Pym in Tarragona inspired me to identify differences between interpreting and written translation. Through my research, I have discussed, as features unique to interpreting, interpreters' proximity and visibility to the parties in a communicative event, the need to address various elements of dynamic interaction and the immediate consequences of their renditions. Given the increasingly diversified ways in which interlingual communication is enabled, it may now be more meaningful to examine "human interpreters v. AI."

Keywords: proximity, visibility, interaction, immediacy, accountability

Introduction: Tarragona, 2004

I met Professor Anthony Pym in 2004 as a student in the doctoral program he had launched in Tarragona a year earlier. I was part of the second cohort of students. The flexibility of the hybrid curriculum comprising online courses and onsite intensive seminars offered an unparalleled opportunity to pursue doctoral study for people like me who already had professional commitments in places where no doctoral program in translation and interpreting studies was available at that time. More importantly, Pym's passion, ingenuity and leadership successfully brought a dream team of translation and interpreting scholars together in this charming little town with Roman ruins. How lucky I was to attend seminars and tutorials given by Pym, Andrew Chesterman, Franz Pöchhacker, Gideon Toury, Ian Mason, Michael Cronin, Miriam Shlesinger and other luminaries! I found the experience at Tarragona highly enlightening and stimulating, but sometimes unsettling in the beginning. As a professionally trained conference interpreter, I felt challenged by unfamiliar notions such as interpreting as an act of mediation, untrained interpreters and *translation* as a term inclusive of interpreting. When Pym asked us in one of the weekly assignments of the online introductory course, "Should there be a separate "map" [as in Holmes' Map of Translation Studies] for interpreting studies as an independent discipline from translation studies?", I must confess

that I boldly declared “yes!” Through exchanges with professors and my own doctoral research, I quickly came to embrace the view that translating texts and interpreting are indeed the same in the sense that they are both social acts of mediating communication across languages and cultures, and I also accepted that conference interpreting is a relatively recent and fractional phenomenon from a historical point of view. My horizons were certainly broadened from the narrow conception I had held of what interpreting was about, i.e., a highly professional act performed by individuals who received specialized training via higher education. Pym likes to say his program “converted” me.

There is one point, however, I did not fully settle with Pym. It concerns the use of the term *translation* itself. Like many scholars, Pym generally treats *translation* in English as an umbrella that encompasses interpreting. For clarification, some scholars (e.g., Daniel Gile and Franz Pöchhacker) use *Translation* (with upper-case T) as an overarching term and *translation* (with lower-case t) for written translation, but, to my knowledge, this distinction is not adopted widely. I remember expressing to Pym my resistance to using *translation* to discuss oral (and signed) rendering, which derived from the acts having dedicated terms, with no superior-subordinate relation, in my native Japanese (*honyaku* for written translation and *tsuyaku* for interpreting; i.e., there is no term that encompasses both written translation and interpreting). Whenever we translate the English word *translation* into Japanese, we first have to figure out if it is referring to written translation only or to both written and oral/signed renditions, then translate it as *honyaku* or *honyaku tsuyaku*, respectively. For instance, I used *honyaku* in translating Pym’s book *Exploring Translation Theories* (2009) (Pym 2010/2020, translated by Takeda) and *honyaku tsuyaku* in translating “Translation and Intercultural Studies”, the name of our doctoral program, since it dealt with interpreting as well.

While it was this technical issue that initially prompted me to question the practical and cross-cultural validity of treating *translation* as inclusive of interpreting, distinct features of interpreting as opposed to written translation have become an underlying interest of my research since then. It has become a habit of mine to ask, “Can this happen in written translation as well?” when I analyze various phenomena involving interpreting. Through my studies of interpreters in the context of war crimes I felt I had developed a stronger basis for arguing the differences between interpreting and written translation. Additionally, recent discussions of whether AI can replace human translators and interpreters have helped me further reflect on the characteristics of interpreting from a different angle.

I am fully aware that there are within translation and interpreting a diverse range of practices which take hybrid forms. For instance, even back in my Tarragona days, I used to interpret in high-stakes depositions (out-of-court testimonies by witnesses) in US jurisdictions, in which realtime English

transcription provided by the court reporter would appear on the computer screen in front of me and I would interpret out of the transcription into Japanese (text-to-speech). In other words, I was doing sight translation when interpreting from English. I was also able to check the transcription of my English rendition being produced out of the note I had taken while listening to the Japanese source utterance. Anecdotal evidence indicates that this is basically how it works in deposition interpreting even today, since the use of Automatic Speech Recognition (ASR) captioning would pose concerns about security and inaccuracies.

Because of the changing practices mainly influenced by the rapid advancement of various translation technologies, it may not be practical or useful now to distinguish and compare “translation” and “interpreting.” Any discussion based on the simplistic notion of “written vs. oral” would be annotated by numerous exceptions. The purpose of the present essay then is not to reinvent the definition of what interpreting is about or what translation is about. (cf., Pöschhacker (2024) proposes a new three-dimensional model featuring agency, embodiment and immediacy to characterize Translation as a graded concept, which would accommodate the use of machine translation tools.) Rather, in gratitude and admiration for Pym’s achievements, I would like to discuss the features of interpreting which I believed over the years were not to be found in text-to-text translation by presenting findings of my research on two topics. The first relates to the issues of interpreters’ physical proximity and visibility to the parties involved in hostilities in the Japanese military context. And the second topic concerns features of interpreting that AI does not seem fully capable of yet by drawing on my research on the use of automated speech translation systems (with spoken input and spoken output) by local governments in Japan.

Proximity and visibility of interpreters in violent conflict

After the end of the Asia-Pacific War (1931–1945), over 100 interpreters in the Japanese military were prosecuted for war crimes and at least 99 of them were convicted at trials administered by different Allied countries in various locations across Asia. At least 32 of these interpreters were executed (see Takeda (2021) for an overview of these trials). In the case of British trials which convicted the greatest number of interpreters (38) of all the nations, the accused interpreters mostly worked for *Kenpeitai* (military police) and in prisoner of war (POW) camps during the war, and they were “concerned” in situations involving ill-treatment of Allied POWs and local civilians in Japanese-occupied territories. An examination of the nature and circumstances of their acts, which have since been deemed war crimes, can highlight the risks interpreters face in violent conflict due to their physical proximity and visibility to their military employer and its confronting party.

Besides the risk of becoming targeted by the opponent as members of the enemy, embedded interpreters may be exposed to risks connected to possible war crimes: the risk of being compelled to physically participate in an unlawful act, the risk of being viewed by the victim as a participant in an unlawful act even for just interpreting and the risk of being called on to testify as eyewitnesses of war crimes after the war (Takeda 2021).

Physical proximity

Unless working from a booth or remotely, interpreters are immediately present with people using their service. In the context of the Japanese military during the Asia-Pacific War, the records of British war crimes trials (see Takeda (2021) for details on the sources) reveal that interpreters mediated communication when their superiors engaged with local residents and POWs in a range of situations, including arrests, interrogation and torture. In some cases, besides interpreting, interpreters were ordered or prompted by their superiors to physically assist torture during interrogations, such as holding the victim's legs or even jointly assaulting the victim. Being alongside their military superiors under volatile and intense conditions meant that interpreters were never exclusively interpreting and had to follow whatever orders their superiors gave. The trial records indicate that interpreters who hesitated to follow orders to participate in violent acts towards civilians and POWs faced severe corporal punishment or threat of court-martial.

Even if interpreters did not engage in assault, victims and witnesses viewed them as participants in the violence by virtue of having been physically present as members of the units that committed such acts. Since the interpreter was always nearby, directly speaking to POWs and local residents in the language they understood, it was easier for these people to remember the interpreter's face and name. Thus, interpreters were prone to being more recognizable than their superiors when war crimes investigators asked the victims and witnesses to identify the perpetrators.

Lastly, interpreters can be called on to testify as eyewitnesses of war crimes if they were present when the crimes took place. At British trials, a number of interpreters gave evidence against defendants who were their former superiors. They were valued as witnesses because of their linguistic understanding of what was happening on both the perpetrator's and the victim's sides. Incidentally, those accused interpreters who also testified for the prosecution were mostly foreign-born Japanese and Taiwanese as well as local hires in Japanese-occupied territories (see Takeda (2021: 140–141) for loyalty-related analysis).

Visibility and “mistaken impression”

The issue of interpreters' visibility in the context of war crimes is closely connected to their proximity to the victims of Japanese atrocities. As discussed above, because interpreters were regularly in direct contact with POWs and local civilians, they were highly visible and likely to make more lasting impressions on the victims of violence than those who actually ordered the violence or executed the orders. For example, Lomax (1995), a former British POW, shares his view of the Japanese interpreter as the centerpiece of his experience being severely tortured. In particular, if interpreters were locally hired or had a similar or familiar background to the victim, they were even more recognizable and became objects of hatred. For instance, Taiwanese mobilized as interpreters because of their ability to speak Japanese and Hokkien were deeply loathed as traitors by Hokkien-speaking ethnic Chinese living in Southeast Asia. Kanao Inouye, a Canadian-born interpreter, was abhorred by POWs, especially those from Canada. That Taiwanese accounted for 10 out of 31 interpreters executed and that Inouye was the only interpreter who was executed without a conviction involving the death of a victim may reflect the amplified visibility of these individuals as “traitors.”

Although interpreters had no rights or authority as civilians to order arrests, interrogation or torture, British trial records contain testimonies and statements by victims and eyewitnesses that accused interpreters of playing the primary or lead role in Japanese wrongdoings. The defense counterargued that it was a “mistaken impression” (Takeda 2021: 13-24) or illusory agency. To the victims, however, the utterances coming from the interpreter led to direct and immediate pain and suffering. It was almost impossible for victims to view interpreters as innocent bystanders who were only delivering the messages of their superiors. To such a victim, the interpreter was fully integrated in the military unit and an enabler of its abuse and cruelties, even if *only interpreting*.

“Proximity of alternative non-linguistic action”

The interpreter's physical proximity to the users of their service is a key feature of non-remote interpreting. Pym (2016: 247) provides a novel angle to understanding its significance by analyzing an interpreter-mediated military encounter in Afghanistan, using the term “proximity of alternative non-linguistic action.” Here, attention is paid to interaction and immediate feedback (including feedback of a physical nature), a feature unique to face-to-face communication. Weighing possible outcomes of how to mediate communication between a village elder and a US sergeant asking him about the Taliban, this interpreter chose to mistranslate in order to avoid the risk of immediate physical harm or other undesirable feedback. This example illustrates how little time the interpreter has to manage the risk of

communication failure, which can lead to his own injury or even death in conflict zones.

Although not addressing extreme cases such as war and violent conflict, Alexieva (1997) also emphasizes direct and active interaction, nonverbal communication and immediate feedback as features of liaison interpreting. She argues that interpreters play a prominent role in making communication successful when there are conflicting goals and asymmetric power relations between the parties, and that in-house interpreters have a greater sense of power compared to freelancers and may behave like principal participants in achieving the goal of the organization they belong to. Embedded interpreters in a military unit may also exercise their agency to make interpreter-mediated communication align with the interests of the unit. Within the military hierarchy and power constellation, however, the scope with which civilian interpreters can exercise their agency is likely to be limited.

On interpreter ethics in war

At British war crimes trials, the two main defenses the accused interpreters resorted to were that they were only interpreting, and that they were following superior orders. Neither of these was effective because the judges sided with the prosecution's assertion of their joint responsibility and the invalidity of the superior order defense when the order was patently unlawful or ethically inhumane. Here, rather than discussing different modes of criminal liabilities for interpreters in war crimes (see Takeda (2021: 127–130) for details), I would like to focus on the case of interpreting torture from the viewpoint of interpreter ethics, drawing on Pym's discussion (2012) of translator ethics. Although the accused were ad hoc interpreters at a time when there were no widely acknowledged guidelines for professional conduct or codes of ethics for interpreters, the topic of interpreting torture from an ethical point of view is relevant even today, as it has continued to take place, with prominent examples from the wars in Afghanistan and Iraq (Senate Select Committee on Intelligence 2014).

Interpreting torture

At British trials the prosecution argued that the accused interpreters were willing to facilitate their units' torture of POWs and local civilians and that they were indifferent to the sufferings of these individuals. Interpreters were even accused of having acted as leaders in some torture cases. In response, the accused interpreters insisted that they were mechanically executing the orders given by their superiors. Whether they were just interpreting or actually participating in physical abuse, the accused interpreters tried to conceal their

agency by emphasizing their “neutral” position “like a machine” (Takeda 2021). The view of interpreters as disinterested machines has been used to help them attempt to evade ethical accountability for problematic acts they enabled through interpreting. For instance, Günter Deckert, who interpreted a Holocaust denier’s speech, defended himself as just a mouthpiece of the source speaker. Pym (2012: 37) calls this “a clever manipulation” of the generally accepted notion that the interpreter is not responsible for the content of what the speaker says.

In his seminal book *On Translator Ethics*, Pym (2012) suggests that translators do not have ethical responsibility for the content of the text since translators are not the author, but translators are responsible for the decision whether to translate a given text and the possible effects of their translation. In the context of interrogational torture by the Japanese military, it could be understood then that these interpreters were not responsible for the content of what their interrogators said, but they were responsible for deciding if they would agree to interpret torture-inducing interrogations and for the potential effects of their interpreting. These effects were immediate indeed, and the victims’ suffering unfolded before the interpreters’ eyes. Whether the interpreter could choose not to interpret is questionable, though, given the doctrine of absolute obedience to superior orders in the Japanese military. However, some interpreters, especially those from Taiwan, noted moral dilemmas and personal disgust in their affidavits and in-court testimonies at British trials.

Under today’s international law, torture and abusive interrogation techniques are illegal and prohibited in all circumstances. Accordingly, interpreters who serve in interrogational torture risk being held criminally liable. As in the case of war crimes trials against the Japanese, the superior order defense would not work since an order to torture is illegal. What transpires in interpreting torture or enhanced interrogation may be an extreme example of how different interpreting is from written translation in the sense of physical proximity to the perpetrator and the victim as well as the immediacy of the effects of cross-lingual exchanges. When the notion of the legality of war or the rules of engagement seem to be ignored by certain parties around the world, interpreters involved in war and violent conflict should be reminded that they can disobey illegal orders (cf., U.S. lawmakers’ call for service members to disobey illegal orders (Jaffs 2025)) and that they could be prosecuted for being involved in manifestly unlawful acts such as the torture and killing of civilians.

Legal consequences aside, the ethics around interpreting torture are articulated in the resolution to condemn torture, adopted in 2007 by the American Translators Association and the National Association of Judiciary Interpreters and Translators in the United States. Against the background of abuse scandals in Iraqi prisons where many interpreters worked, the resolution states that “knowing participation in, facilitation or countenancing of,

cooperation with, or failure to report torture or other mental or physical abuse or degradation of any human being” (ATA/NAJIT 2007) violates the code of ethics. This resolution should be revisited in the discussion of interpreter ethics in training and professional settings from the standpoint of exceptions to professional secrecy as well (see Takeda (2021: 145-149) for detailed discussion).

AI clarifying the work of human interpreters

Beyond the context of interpreters in war and violent conflict, there is another development that has made me think further about characteristics unique to interpreting: the rapid advancement of AI-based technology related to interpreting. This situation has been challenging interpreting scholars to reconsider what is involved in the process of interpreting as well as how future interpreters should be trained. I myself have been asked every so often about the future viability of interpreting as a profession. In response, a decade ago, I took advantage of the invitation to contribute a short Japanese-language essay to the journal of an association of machine translation researchers in Japan. Here, I first present a summary of this small piece, the title of which translates to “Will AI replace interpreters?”, and examine how my view of AI and interpreting may have changed over the last ten years, mainly by drawing on the findings of my research on the use of automated speech translation systems by local governments in Japan.

Will AI replace interpreters? (2016)

What follows is an English summary of the essay I wrote in Japanese on AI and interpreting in 2016 (Takeda 2016).

Communication in spoken language is a dynamic process jointly constructed by participating speakers through their interactions, which are spontaneous and immediate. There are paralinguistic elements such as prosody, nonverbal elements like gestures, and temporal elements including pauses. The meaning of speech is formed through the interplay of all these elements. Further, spoken language involves colloquial expressions, accents, dialects, honorifics and nuances that differ from written language. These issues pose challenges to automated speech translation. Above all, there are ethical concerns as to who would be held accountable for the quality of the renditions. Who would be held liable for errors that could lead to serious consequences? There is no step for humans to correct errors of AI-generated interpretation, like post-editing of translation.

Prior research has demonstrated that interpreters do more than interpret. For instance, they also manage turn-taking among speakers

and intervene to resolve communication glitches. Interpreters must also pay attention to the power dynamics among the speakers and the characteristics of the audience, selecting appropriate honorifics and registers. To what extent can automated speech translation address these pragmatic aspects of interpreting?

AI-based interpreting has great potential to be useful for non-critical, short, transaction-based communication. If it supports less-taught languages, it would be socially beneficial. Interpreters could possibly be replaced by automated speech translation if speakers read out prepared scripts faithfully, and audiences are prepared to endure a monotonous synthetic voice for an extended time. The use of AI interpreting in judicial settings would not be appealing due to the complicated issue of evidentiary admissibility and accountability.

A decade later, some of the questions I raised may have been addressed. Almost every day, Google alerts me to news on the latest improved features and the increasing use of “AI-powered speech translation” reported by various developers and technology writers. According to them, there is now technology that can provide realtime speech translation, while preserving the source speaker’s tone, emotions and vocal characteristics in the output (Stasimioti 2025), and the use of AI-powered speech translation has been reported in a wider range of settings, including tourism, customer service and dubbing (Certified Languages International *n.d.*). Also, more interpreters seem to be embracing the practice of AI-assisted or AI-supported interpreting (see various announcements of seminars on AI tools for interpreting on the AIIC website, for instance). Incidentally, the AIIC Science Hub released an “AI Interpretation Checklist” in 2025 as a tool for interpreters to use when advising clients whether to use “AI Interpretation (or more precisely, Automated Speech Translation–AST)” (AIIC 2025). Although it is meant to suggest in what context AST may be a feasible option, the list succinctly reveals limitations of AST, which resonate with some of my concerns from a decade ago – namely, the difficulty of handling interactive exchanges between speakers and the issue of accountability. Against the backdrop of this rapid development in AI-based speech translation and the debate on its limitations, I now present an example of how such speech translation systems are used in real life (i.e., discourse that occurs organically, rather than as an experiment or for demonstration purposes).

Government use of automated speech systems for foreign residents in Japan

Collaborating with a colleague in Public Administration, I conducted a survey in 2023 with 115 local governments (hereafter, LGs) and follow-up interviews with 26 LGs in Japan on their use of automated speech translation systems in assisting Ukrainian “evacuees” (the term the Japanese government has

adopted to circumvent highly restrictive refugee policies). For the discussion of this particular case, I also adopt the Japanese government's language of "foreign nationals" and "foreign residents" instead of preferred designations such as "persons with limited Japanese proficiency" or "linguistically and culturally diverse communities." With the increasing number of foreign nationals living in Japan, many LGs have been providing translation and interpreting services in English, Chinese and Korean in various situations, including registrations and disaster preparedness. Some LGs are also staffed with speakers of Portuguese, Spanish, Filipino and Vietnamese. Ukrainian, however, was a new language for these LGs, with very few people available who could provide language assistance. Some Ukrainians did not wish to face Russian nationals as interpreters or speak Russian themselves even when capable. Incidentally, the Japanese central government had been funding programs to develop automated speech translation systems and promoting their use among LGs. Thus, the aim of the survey and interviews was to find out to what extent automated speech translation systems were used by LGs in communicating with Ukrainian evacuees.

The results of the survey and interviews (Takeda and Inagaki 2023) indicated that about 57% of the respondents had used automated speech translation systems mainly because they were free of charge and easy to use. Among that group a little over 50% reported that the system always or almost always served the purpose, but one third shared various problems they encountered: the systems did not work very well with long sentences, complex or culture-based content, administrative terms, proper nouns, idioms and different accents; it was difficult to have smooth communication because of the time lag; and the system struggled to "understand" the input, necessitating speaking in short sentences or repeating oneself. Only two respondents felt assured about security issues. Those who never or hardly used automated speech translation systems pointed to concerns about privacy protection.

Since 2025 we have been conducting interviews with other LGs and their affiliated organizations on how they operate "consultation systems for foreign nationals", focusing on language issues. One of the inquiries concerns the potential use of the "AI simultaneous interpreting (SI) systems", which were recently developed by a government-funded project and promoted by the government as a solution to the language barriers in public settings where foreign nationals are involved (Japanese Ministerial Council on the Acceptance and Coexistence of Foreign Nationals 2025a, 2025b). None of the four organizations we have interviewed so far reported any use of AI-based SI systems at this time (Takeda and Inagaki 2025).

Dehumanized communication in consultation

There is a brochure issued in 2021 by the Advanced Information Systems and Software Division of the Japanese Ministry of Internal Affairs and Telecommunication to promote the use of automated speech translation systems among LGs. Incidentally, this is the same ministry (albeit a different division) that has funded programs to develop MT/AI-based speech translation systems. The brochure claims that one of the benefits of utilizing automated speech translation systems is that they can ease the psychological burden experienced by some foreign residents when seeking consultation on sensitive issues such as domestic violence and welfare. According to the brochure, some such foreigners prefer machines over human interpreters in these situations because they feel uncomfortable if the interpreter comes from their own community. Users' discomfort or distrust towards interpreters has been reported in the context of asylum seekers (e.g., Pokorn and Čibej 2018). In civic service settings the brochure focuses on, however, remote interpreting is a possible solution. The problem with AI-based speech translation systems here is that without a clear understanding of how data is protected and managed as well as the client's consent, it becomes a concern to let sensitive personal information be fed into the systems. Vieira (2026: 79-80) also points to the danger of using MT in an ad hoc manner without being aware of the issue of privacy protection. Besides, a speaker's struggle with talking about complex problems coherently in times of distress can pose a great challenge for AI, especially with low-resource languages, which are often used in community settings. This can result in translation inaccuracies, which may in turn lead to additional undesirable consequences.

As mentioned earlier, interpreting is a complex task that requires constant attention to paralinguistic and nonverbal elements in communication as well as power dynamics among the speakers and their cultural backgrounds. Interpreters exercise their agency, when necessary, for clarification or repair of problems that may arise in the course of spontaneous and immediate exchanges. I am uncertain if AI could ever address every aspect that human interpreters do. Attention should also be paid to how interpreters manage their positionality and account for the choices they make in interpreting (cf. Martin 2025). As in the case of AI and MT in general, how the issue of accountability in AI-based speech translation is settled remains to be seen.

Conclusion: Shift to “human v. machine”?

Looking back, I realize now that I have hardly considered linguistic forms or surface-level characteristics for differences between written translation and interpreting, such as the use of the plain form and the polite form in Japanese.

(The polite form is almost always used in interpreting, while both the plain form and the polite form are used in written translation.) I do not mean that linguistics-focused research is irrelevant or unimportant. As a practicing interpreter, however, I have been keenly aware that interpreting is a socially situated practice. Thus, my attention has always been directed to how interpreter behavior is influenced by social and situational factors, including power relations, visibility and institutional constraints. I have never been formally trained in sociology, but some people tell me that I take a sociology-like approach. I suppose I have indeed been “converted” by the Tarragona program.

In this essay I have discussed some features I believed unique to interpreting, including interpreters’ physical proximity and visibility to the parties in a communicative event, their need to handle dynamic exchanges with attention to power relations and to exercise their agency, if necessary, to prevent or fix communication failures, and the immediate consequences of their renditions. Given the increasingly diversified ways in which cross-lingual communication is enabled, I cannot say with absolute certainty that these attributes cannot be found in written translation or translator behavior. We might observe instances, however rare they may be, where translators work on text alongside their clients who need quick output and the translators revise their work immediately in response to their feedback.

Then, can translators who render cross-lingual communication output in text be exposed to the risk of prosecution for war crimes such as killing civilians and torture? Maybe not, because their work would not constitute a direct element of the criminal act. On the other hand, interpreters could be prosecuted for interpreting if they are present and their renditions are directly involved in interrogational torture. It should be noted, however, that translators have also been persecuted, tortured or even killed for translating certain texts (such as the purported *fatwa* killing of Hitoshi Igarashi, who translated *The Satanic Verses* into Japanese (Weisman 1991)).

Ultimately, as suggested earlier, it may not be a productive endeavor to seek unique aspects of interpreting as opposed to written translation at this time. My interest in pursuing differences between interpreting and written translation was certainly inspired by my conversations with Pym in Tarragona. That pursuit now seems to be evolving into the question of what aspects of interpreting will probably not be addressable by AI, which should help me clarify what to focus on in my research into how humans communicate across different languages and cultures. I am hoping I can still pick Professor Pym’s brain throughout my efforts. Now, what would happen if AI-based speech translation were used in interrogational torture?

Declaration of AI use

The author did not use any AI-assisted tools in the preparation of this manuscript.

References

- Advanced Information Systems and Software Division of the Japanese Ministry of Internal Affairs and Telecommunication. 2021. “Guide to implementing ‘multilingual speech translation service’ at local governments” (in Japanese). https://www.soumu.go.jp/main_content/000745491.pdf. Visited May 2026.
- AIIC Science Hub. 2025. *AI Interpretation Checklist*. <https://aiic.org/company/roster/companyRosterDetails.html?companyId=14022&companyRosterId=120>. Visited May 2026.
- Alexieva, Bistra. 1997. “A typology of interpreter-mediated events”. *The Translator* 3(2): 153-174.
- American Translators Association. 2007. “Resolution of American Translators Association”. *ATA Chronicle*, November/December 2007: 11.
- Certified Languages International. *n.d.* “AI Speech Translation”. <https://certifiedlanguages.com/blog/ai-speech-translation/>. Visited May 2026.
- Jaffe, Greg. 2025. “Democratic lawmakers tell military to refuse illegal orders”. *The New York Times*, 18 November 2025. <https://www.nytimes.com/2025/11/18/us/politics/democratic-lawmakers-illegal-orders.html>.
- Japanese Ministerial Council on the Acceptance and Coexistence of Foreign Nationals. 2025a. “Comprehensive Measures for Acceptance and Coexistence of Foreign Nationals (revised in 2025)”. https://www.moj.go.jp/isa/support/coexistence/nyuukokuKanri01_00140.html. Visited May 2026.
- Japanese Ministerial Council on the Acceptance and Coexistence of Foreign Nationals. 2025b. “Roadmap for the Realization of a Society with Harmonious Coexistence with Foreign Nationals (revised in 2025)”. https://www.moj.go.jp/isa/support/coexistence/04_00033.html. Visited May 2026.
- Martin, Conor. 2025. “Understanding the complexity of interpreter positionality of interpreters and positionality management in armed conflict and beyond”. In M. Todorova and L. Ruiz Rosend (eds.) *The Routledge Handbook of Translating and Interpreting Conflict*. London: Routledge. 198-220.

- National Association of Judiciary Interpreters and Translators. 2007. "Resolution Condemning and Deploring Torture adopted by the membership of the National Association of Judiciary Interpreters and Translators at the 2007 NAJIT Annual Meeting May 19, 2007 – Portland, Oregon". *Proteus*, Fall 2007: 16.
- Pokorn, Nike. K. and Čibej, Jaka. 2018. "Interpreting and linguistic inclusions – Friends or foes?". *The Translator*, 24(2): 111-127.
- Pöchhacker, Franz. 2024. "Is machine interpreting interpreting?". *Translation Spaces*. <https://doi.org/10.1075/ts.23028.poc>
- Pym, Anthony. 2010/2020. *Honyaku riron no tankyu* [Exploring Translation Theories]. Translated by Kayoko Takeda. Tokyo: Misuzu Shobo.
- Pym, Anthony. 2012. *On Translator Ethics: Principles for Mediation Between Cultures*. Amsterdam and Philadelphia: John Benjamins.
- Pym, Anthony. 2016. "Risk analysis as a heuristic tool in the historiography of interpreters: For an understanding of worst practices". In K. Takeda and J. Baigorri-Jalón (eds.) *New Insights in the History of Interpreting*. Amsterdam: John Benjamins. 247-268.
- Senate Select Committee on Intelligence. 2014. *The Senate Intelligence Committee Report on Torture*. Brooklyn, New York: Melville House.
- Stasimioti, Maria. 2025. "Uni and startup explore how to preserve voice, emotion in AI speech translation". *Slator*, 16 October 2025. <https://slator.com/uni-startup-explore-how-preserve-voice-emotion-ai-speech-translation/>. Visited May 2026.
- Takeda, Kayoko. 2016. "Will AI replace interpreters?" (in Japanese). *The AAMT Journal*, 67:1-2.
- Takeda, Kayoko. 2021. *Interpreters and War Crimes*. Milton Park, Abingdon, Oxon: Routledge.
- Takeda, Kayoko and Inagaki, Hiroshi. 2023. "Local governments' use of speech machine translation with Ukrainian evacuees in Japan" (in Japanese). *Interpreting and Translation Studies*, 23: 37-59.
- Takeda, Kayoko and Inagaki, Hiroshi. 2025. "Multilingual consultation systems in Japan: Local governments' practices and the central government's 'measures on foreign nationals'" (in Japanese). *Interpreting and Translation Studies*, 25: 83-102.
- Vieira, Lucas Nunes. 2026. *Translate: Multilingual Access to Critical Services in the Age of AI*. Cambridge, UK: Cambridge University Press.
- Weisman, Steven R. 1991. "Japanese translator of Rushdie book found slain". *The New York Times*. 13 July 1991. ProQuest.

This volume presents texts in honor of Professor Anthony Pym.

Introduction, editorial team

Prologue, Andrew Chesterman, *Anthony Pym, on borders*

Nataša Pavlović, *The relevance of translation theory for translator education today*

Volga Yilmaz-Gümüş, *Tracing the evolution of translator training in Anthony Pym's research*

Judith Raigal-Aran and Nune Ayvazyan, *Watching the process unfold: using API-based interaction data to understand student use of AI in translation tasks*

Maggie Hui, *A constraints-first approach to CAT and generative AI in translator education: a simulated client project in a master's course*

Enes Ekici And Şeyda Eraslan, *Beyond perceptual failure: cognitive load and number renditions in technology-mediated English–Turkish simultaneous interpreting*

Yu Hao, *Quality in multilingual crisis communication: a reception-oriented perspective*

Juerong Qiu, *How does multimodal creativity travel? A reception perspective on cyberdubbing*

Gabriel González Núñez, *“Art, esthetics, ideal”: an ethics for literary translators in the age of translating machines*

Snježana Veselica Majhut, *How to find Paula D.? In quest of translators' identities in early 20th century Croatia*

Kayoko Takeda, *On interpreting v. translation*

