Research Article

Applied Linguistics

THE NEED FOR COMPARABLE CORPORA IN TRANSLATION TEACHING AND LEARNING

Iris Klosi	University of Tirana, Faculty of Foreign Languages, English Department
Ergys Bezhani	University of Tirana, Faculty of Foreign Languages, English Department

Abstract

      The primary objective of this research is to catalyze the scholarly progress of translation teaching and learning within higher education institutions in Albania, specifically in the preparation of students in the field of translation theory and practice. The study aims to contribute significantly to the theory and practice of translation by emphasizing the imperative to develop comparable corpora of English-Albanian texts across diverse text types. Additionally, the research will delve into establishing a typology of translation strategies in Albania. Building upon the pioneering efforts of the researchers in previous work, which involved the translation of English and Albanian texts using linguistic, stylistic, and semantic techniques, this study also pays meticulous attention to the intricate relationship between the translation process and its resulting product. By focusing on the development of comparable corpora, this research presents a valuable opportunity to enhance both the theory and practice of utilizing bilingual corpora in the context of translation teaching, learning, and research. The outcomes of this research endeavor will provide the foundational framework for the future formulation of a streamlined pedagogical typology specifically addressing the use of comparable corpora in teaching and learning translation strategies in the English-Albanian context. This typology, once developed, will further serve as the basis for generating a series of project concepts with the overarching goal of establishing cooperative agreements. These agreements are envisioned to foster collaboration in the future with colleagues from regional and Western universities, as well as with researchers and experts in the field of translation.

Introduction

The current landscape of research in corpus-based approaches within the realm of translation theory and practice is relatively novel in Albania. A handful of privately-owned companies specializing in translation, conference interpreting, and localization services are endeavoring to usher in the latest trends in computer-assisted translations, particularly leveraging translation memory tools. The unfolding dynamics of societal advancements in Albania, spurred by reforms across various facets of human life and the ongoing EU accession process, have compelled practitioners, researchers, and experts to prepare for the alignment, transposition, and approximation of EU legislation. This extends to ensuring compliance of national laws with the EU acquis across diverse domains.

The initiation of a process to unify terminology across different realms of human endeavor dates back to 2002 when the former Ministry of Integration in Albania called upon official translators, including many university lecturers, to translate and standardize terminology. The objective was to establish consistency and accuracy throughout the translated documents. Despite this concerted effort, the vast repository of standardized terms and equivalences, painstakingly compiled by professionals and official translators licensed by the Ministry of Justice, largely remained stored within the Ministry's computers, without subsequent exploration or utilization by practitioners.

Notably, a private-owned company in Tirana, Albania, has recently introduced the SDL Trados Studio as a computer-assisted translation (CAT) tool and translation memory software. This company holds exclusive rights to resell it to freelancers, partners, and competitors, and also conducts training sessions for interested individuals or groups. During the Erasmus+ Teaching Mobility Week in April 2022, two professors from the University of Applied Sciences in Nysa, Poland, introduced MemoQ to faculty colleagues, undergraduates, and postgraduate students. Preliminary discussions about a potential joint project on CAT tools were initiated during this event.

A newly enrolled PhD student at the English Department, beginning in September 2022, has expressed interest in computer-aided applied linguistics. This student focuses on the automated treatment of language for specific purposes, with guidance from a colleague in the French department who has expertise gained through participation in the A Corpus-Based Method for Extraction and Polylexical Units Elexis Project, funded by Horizon 2020 in Copenhagen, Denmark. Another colleague has pursued postgraduate studies in Linguistics and Web Technology at the University of Marburg, Germany. However, there is currently no review of the existing curricula to introduce corpora-based translation as a separate course or research field at the university.

Despite the burgeoning interest among young scholars and professionals, the scarcity of expertise exacerbates the need for increased cooperation in the field of corpus-based and corpus-driven translation. Moreover, the absence of studies and research in neighboring countries such as North Macedonia and Kosovo, where Albanian is an official language, underscores the necessity for collaboration and networking transcending state boundaries.

Theoretical framework

The current research landscape, which is predominantly concentrated in the Western European and US regions, is prominently directing its focus towards the integration of new technologies in audio-visual translation, localization, globalization, and corpus-based approaches (Munday et al, 2022, pp.299). Each of these facets has triggered a profound re-evaluation of both translation theory and practice (Munday et al, 2022, pp.299). The adoption of corpus-based, corpus-driven, machine translation, localization, and globalization strategies not only entails studies in translation but also delves into the behavioral and cognitive effects on practitioners, scholars, researchers, and society at large. In bilingual societies, where the demand for real-time document translation is substantial, there is an increasing reliance on localization or machine translation options, driven by the desire to save both time and resources.

Internationally, scholars and researchers in the realm of corpus-based translation advocate for the exploration of the immense potential offered by large computerized corpora. A notable shift in contemporary translation studies is observable, wherein source texts are no longer perceived as the exclusive starting point for translation analysis. Instead, target texts assume a pivotal role in elucidating and contrasting the norms and rules inherent in the source texts. As aptly expressed by Baker (1993, pp.248), "The status of the source texts has been undermined, and we have managed to make a leap from source-text-bound rules and imperatives to descriptive

Page | 36

categories." She argues for the necessity of a robust research methodology and a set of adaptable tools to effectively employ comparable bilingual corpora and develop enduring and effective translation strategies and methodologies.

Munday et al. (2022), categorizes corpora into three types: monolingual, comparable bilingual, and parallel corpora. Monolingual corpora consist of texts in the same language, serving researchers in identifying characteristics of genre or author style, as well as aiding practitioners in assessing naturalness, including frequent collocations. Large monolingual corpora, such as the British National Corpus and the COBUILD Bank of English, function as representative reference corpora, offering a benchmark against which language deviations can be measured.

Comparable bilingual corpora, according to Munday et al. (2022), are specialized collections of similar source texts (STs) in two languages that are ripe for the exploration and extraction of terminology and equivalences. However, the definition of comparable corpora varies. Delpech (2014, p.7) presents a narrow definition, encompassing texts in two or more languages complying with the same genre, field, and sampling period criteria, maintaining balance and representativeness. A broader definition necessitates a substantial subpart of vocabulary in common between texts (Delpech, 2014, p.7). Delpech adopts a middle ground, considering sets of texts as comparable if they are in two or more languages, dealing with the same topic, and preferably generated within the same communication situation (2014, p.7). Corpora can also be monolingual, comprising translational and non-translational sub-corpora, sharing genre, proportions, publication time frame, and domains (Diriker, 2014, p.60).

Parallel corpora, consisting of source text (ST) and target text (TT) pairs, allow for the aligned investigation of strategies employed by translators when examined sentence by sentence or paragraph by paragraph (Kenny, 2001, 2011). Examples of parallel corpora include Linguee, MyMemory, OPUS, and the Canadian Hansard.

The impetus for developing more sustainable strategies in translation theory and practice draws heavily from corpus-based linguistics, particularly when exploring grammatical features of written or spoken languages. The need to provide a robust basis for data collection emerges as a critical concern in the realm of corpus linguistics (McEnery and Hardie, 2012, p.6).

Research Questions

Corpus-based translation and research have garnered limited academic interest in Albania, reflecting a broader expertise not only within the country but also across the Balkan region. Consequently, noticeable gap exists in the development of a typology concerning corpus-based translation strategies in English-Albanian, with a specific focus on the concept within a linguistic and cultural context. The underexplored nature of corpus-based translation in this region presents numerous challenges and obstacles to its effective implementation.

The majority of master's and doctoral theses in the field of agro-translation in Albania have traditionally concentrated on theoretical and practical aspects across various domains, such as literature, drama, terminology, and poetry translation. However, these studies predominantly adopt a descriptive approach to translation analysis or criticism, primarily centered on translation dichotomy and analysis from the sourcetext. Recognizing a growing interest in comprehending

and analyzing translated products, there is an emerging need to develop a robust research methodology to support practitioners, academics, students, and researchers involved in corpus-based translation. Currently, there is a notable absence of corpora, repositories, inventories, or machinetranslation systems established within the Albanian context. Furthermore, contemporary researchers at the Faculty of Foreign Languages, University of Tirana, Albania, have continued to delve into the use of copyright-based translation as a multifaceted tool for teaching, learning, research, and practical applications.

Freelance translators and interpreters collaborate in various domains to align terminologies, employ accurate syntactic patterns, and convey content effectively. However, their focus is more on practical aspects, lacking awareness of the essential theoretical rules and guidelines that could enhance the success and appreciation of the work during the transition from sourcelanguage to the target language. Additionally, there is a lack of recognition among these professionals regarding the nuanced translation of stockmetaphors, idioms, and expressions found in different contexts, emphasizing the need for a deeper understanding of when identified and mined within comparable bilingual corporafromsimilar sources, targets, or original texts.

The urgency for the development of a series of project concepts aimed at fostering collaboration between researchers and scholars from the Albanian universities and the tribounterparts of the Balkans and Eastern and Western European countries is evident. The absence of opportunities or networks among practitioners, researchers, and students in Albania further supports the challenges in the field of corpus-based translation, highlighting the pressing need for comprehensive solutions and adenylesfordialogue and exchange.

Methodology of Research

The research methodology employed in this study integrates both theoretical underpinnings and practical applications, focusing on the exploration of comparable bilingual translation. Central to this approach is the emphasis on recognizing and addressing the imperative need for the development of comparable bilingual corpora tailored for the enhancement of translation teaching and training endeavors.

The foundational principle of the methodology lies in laying the groundwork for tangible and domain-specific corpora that can effectively assist aspiring professionals in the realms of translation theory and practice. This involves a meticulous process of preparing a robust foundation to cater to the diverse needs of future practitioners by creating corpora that are not only comprehensive but also contextually relevant across various fields of life.

To achieve these objectives, the methodology places a significant emphasis on the collection and analysis of qualitative data sourced from a myriad of channels. This inclusive approach involves tapping into multiple open sources, both online and offline, encompassing a broad spectrum of materials in both English and Albanian languages. The selection of materials prioritizes similarity without necessitating direct parallel equivalence, allowing for a nuanced exploration of translation dynamics.

The dataset for analysis comprises a diverse array of components, including source texts, formal and informal translations of these source texts, and original texts in Albanian. This meticulous selection process is instrumental in capturing the intricacies and variations inherent in translation practices. By leveraging this diverse array of data sources, the research aims to provide a comprehensive understanding of the challenges and opportunities that arise in the translation process.

In essence, the methodology strives to bridge the gap between theoretical knowledge and practical application by fostering the creation of tailored corpora. This not only facilitates effective translation teaching and training but also ensures that future professionals are equipped with the tools and resources necessary to navigate the intricacies of translation theory and practice across diverse domains of knowledge and application.

Findings

The use of comparable corpora emerges as a cornerstone in the translation classifier, providing students with a practical tool for navigating and understanding the intricacies of translation challenges. By exposing students to diverse sources and text types in both English and Albanian, the study underscores the importance of offsetting an environment where the yycan analyzes, compares, and contrasts various linguisticelements. The richness of comparable corpora allows students to delve into the layers of meaning, styles, registers, and contexts inherent in different discourses.

One key finding of the research is the versatility of comparable corpora derived from diverse sources and text types in two languages. The nuanced shifts in words and phrases from sourcelanguage to the target language became apparent, prompting students to engage in a Thoroughexaminationoftranslationequivalents. This study emphasizes the need for students to observe and interpret meanings, establish translation equivalents, identify styles and registers, and make informed inferences about both denotative and connotative meanings.

Moreover, the research highlights the importance of a dynamicequivalence approach, encouraging students to produce new texts by rearranging and restructuring lexical and cultural units. This approach aims to move beyond traditional textual equivalence, fostering a more dynamic understanding that considers cultural nuances and linguistic richness. The functional method in translation, focusing oncontextually motivated keywords and recurrentphrases, has emerged as a valuable aspect for students to grasp subtleties of flanguage use.

In addition to exploring linguistic skills, researchers have endeavored to delve into mock metaphors, figurative language, mood of expression, and discourse. This broader exploration enhances students' language proficiency, enabling them to expand their vocabulary beyond the confines of dictionaries and grammar. The studies suggest that the mechanism of this multifaceted approach not only aids students in their current academic pursuits but also opposes them for the real-world challenges of professional translation.

Conclusions

This research sets the stage for the development of comparable bilingual corpora, positioning them as essential tools for advancing both theory and practice. The envisagedcorpora, with its own resources and text types, serve as repositories of linguistic richness and cultural nuances. By providing a structured framework for analyzing and contrasting concepts, ideas, and communities across languages, comparable corpora have emerged as catalysts for enhanced translation methodologies.

The development of comparable bilingual corpora in English and Albanian was anticipated to foster collaboration among professionals in the field. This study identifies a future where translators and cholars can collectively explain the latest splice residues beyond the limitations of traditional dictionaries. This collaborative effort aims to enrich the understanding of how words and phrases are used in context, their frequency in various texts, and the nuances of meaning encapsulated in different discourses.

Furthermore, the paper calls for attention from scholars in the Western Balkans, urging them to consider cooperation for the development of comparable corpora. The shared commitment to the project concept and the collective exploration of comparable cooperation, training, and research purposes are seen as crucial steps toward elevating translation studies in the region. This study revealed a network of practitioners, researchers, and students collaborating to bridge the gap between theoretical knowledge and practical application in the field of corpus-based translation, ultimately contributing to the advancement of translation studies in Albania and beyond.

References

Baker, M. (2019). Corpus Linguistics and Translation Studies*: Implications and applications. In Researching translation in the age of technology and global conflict (pp. 9-24). Routledge. https://doi.org/10.4324/9780429024221

Diriker, E. 2014. Corpora. In Routledge encyclopedia of translation studies, Second edition, eds. M. Baker and G. Saldanha. London and New York: Routledge. https://doi.org/10.4324/9781315678627

Delpech, E. M. (2014). Leveraging Comparable Corpora for Computer-assisted Translation. Comparable Corpora and Computer-Assisted Translation, 1-39. https://doi.org/10.1002/9781119002659.ch1

Baker, M., Francis, G. (Ed.), & Tognini-Bonelli, E. (Ed.) (1993). Text and Technology: In Honour of John Sinclair. John Benjamins Publishing Company. https://doi.org/10.1075/z.64

Munday, J., Ramos Pinto, S., & Blakesley, J. (2022). Introducing Translation Studies: Theories and Applications (5th ed.). Routledge. https://doi.org/10.4324/9780429352461

McEnery, T., & Hardie, A. (2011). Corpus linguistics: Method, theory and practice. Cambridge University Press. https://doi.org/10.1017/CBO9780511981395

Zanettin, F. (1998). Bilingual comparable corpora and the training of translators. Meta, 43(4), 616-630. https://doi.org/10.7202/004638ar