Artur Dariusz Kubacki / Piotr Sulikowski (Hg.)

Aktuelle Trends in der Übersetzungswissenschaft

V&R unipress

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

Translation Landscapes – Internationale Schriften zur Übersetzungswissenschaft

Band 8

Herausgegeben von

Artur Dariusz Kubacki und Piotr Sulikowski

Translation Landscapes

Die Bände dieser Reihe sind peer-reviewed.

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

Artur Dariusz Kubacki / Piotr Sulikowski (Hg.)

Aktuelle Trends in der Übersetzungswissenschaft

Mit 35 Abbildungen

V&R unipress

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

Uniwersytet Komisji Edukacji Narodowej w Krakowie

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über https://dnb.de abrufbar.

Diese Publikation wurde von der Universität der Kommission für Nationale Bildung Kraków und der Universität Szczecin finanziert.

Das Buch ist das Ergebnis der wissenschaftlichen Zusammenarbeit im Bereich der Forschungen zur Translationswissenschaft zwischen dem Lehrstuhl für Deutsche Sprachwissenschaft der Universität der Kommission für Nationale Bildung Kraków und dem Institut für Sprachwissenschaft der Universität Szczecin.

Gutachter: Univ.-Prof. Dr. habil. Paweł Bąk (Universität Rzeszów) und Dr. habil. Marcin Walczyński (Universität Wrocław)

Druck und Bindung: CPI books GmbH, Birkstraße 10, D-25917 Leck
Printed in the EU.

Vandenhoeck & Ruprecht Verlage | www.vandenhoeck-ruprecht-verlage.com

ISSN 2747-4496
ISBN 978-3-8470-1754-7

Inhalt

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

Vorwort

Wir freuen uns sehr, Ihnen den achten Konferenzband der Verlagsreihe *Translation Landscapes – Internationale Schriften zur Übersetzungswissenschaft* (TraLa 8), die seit 2023 im Verlag V&R unipress der deutschen Verlagsgruppe Brill Deutschland erscheint, herausgegeben von Artur Dariusz Kubacki und Piotr Sulikowski, vorstellen zu können. Der vorliegende Band ist das Ergebnis der 7. Internationalen Übersetzungswissenschaftlichen Konferenz, die von der Universität der Kommission für Nationale Bildung Kraków und der Universität Szczecin veranstaltet wurde, d. h. von zwei, seit 2020 auf wissenschaftlicher Ebene dauerhaft miteinander kooperierenden Universitäten.

Die oben erwähnte wissenschaftliche Konferenz fand vom 26. bis zum 28. Oktober 2023 in Pobierowo bei Szczecin statt und wurde von mehr als 50 Teilnehmern:innen besucht, von denen 37 Übersetzungswissenschaftler:innen aus Polen, Deutschland, Lettland und der Ukraine ihre Vorträge im Plenum präsentierten. Polen war mit Lehrenden aus Katowice, Kraków, Lublin, Opole, Poznań, Szczecin, Warszawa und Wrocław stark vertreten. Die Konferenzsprachen waren Englisch, Deutsch und Polnisch.

Band 8 trägt den Titel *Aktuelle Trends in der Übersetzungswissenschaft* und stellt eine Sammlung von Texten dar, die sich mit den Ergebnissen der Translatorik und aktuellen Forschungstrends befassen, die in der polnischen und ausländischen Übersetzungswissenschaft heutzutage vorherrschen. Die Themen der Beiträge sind vielschichtig und reichen von der literarischen Übersetzung und Fachübersetzung über das Sprachmitteln bis hin zu modernen Translationstechnologien.

Wie bereits im letzten Jahr wurde auch der diesjährige Konferenzband in zwei Themenbereiche gegliedert: Der erste umfasst das Fachübersetzen, der zweite konzentriert sich auf das literarische Übersetzen. Insgesamt sind 26 Texte in deutscher und englischer Sprache enthalten, von denen jeweils 13 Aufsätze in die beiden vorgenannten Bereiche eingestuft werden können. Der Band schließt – traditionsgemäß – mit Kurzvita zu den einzelnen Autoren:innen in deutscher Sprache ab.

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

Wir bedanken uns sehr bei der Leitung der Universität der Kommission für Nationale Bildung Kraków und der Universität Szczecin für die großzügige Finanzierung dieser Publikation. Unsere Dankesworte gehen auch an die Gutachter des Bandes, Univ.-Prof. Dr. habil. Paweł Bąk von der Universität Rzeszów und Dr. habil. Marcin Walczyński von der Universität Wrocław, für ihre konstruktive Kritik bei der Erstellung dieser Publikation und die wertvollen sprachlichen Hinweise zu den einzelnen Beiträgen.

Artur Dariusz Kubacki und *Piotr Sulikowski*

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

Dolmetschen – Übersetzen – Sprachmitteln – Translationstechnologie

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

Bolesław Cieślik (Uniwersytet Komisji Edukacji Narodowej w Krakowie, Kraków) / Bożena Iwanowska (Akademia Ekonomiczno-Humanistyczna w Warszawie, Warszawa) / Yan Kapranov (Akademia Ekonomiczno-Humanistyczna w Warszawie, Warszawa)

Cluster-corpus translation method in action: verifying legal translatemes in English, German, and Polish corpora

Abstract
The article explores the application of the cluster-corpus translation method using the JRC-Acquis Sub-Corpus of OPUS. It presents a systematic framework for legal translation, focusing on classifying, analyzing, and verifying legal translatemes – the smallest units of translation – in three languages. The paper highlights the method's effectiveness in ensuring linguistic accuracy and contextual appropriateness, particularly in the intricate domain of legal translation. The authors demonstrate this through detailed analysis of specific legal terms and phrases, showcasing the method's utility in handling complex legal terminology across different legal systems and languages.
Keywords: cluster-corpus translation method, JRC-Acquis Sub-Corpus, OPUS, legal translation, translateme

Introduction

Translating legal texts plays a pivotal role in international law and commerce, confronting specific challenges at both lexical and syntactic levels. Scholars in translation studies, such as Chan (2004) and House (2015), emphasize the importance of *translational linguistics.* This critical branch within translation studies bridges the gap between theoretical understanding and practical translation application. It is deeply interwoven with linguistic theories, literary analysis, and philosophical inquiries, aiming to dissect and comprehend the multifaceted nature of translation processes. Translational linguistics investigates how linguistic, cultural, and contextual elements influence translation outcomes, thereby holding a crucial place in translation studies. It delves into essential concepts like equivalence, adaptability, and the translator's role alongside methodologies that enhance translation quality. By doing so, translational linguistics provides invaluable insights into the mechanics of translation, making it indispensable for tackling the complexities encountered in translating legal and other specialized texts. This discipline enriches our understanding of the theo-

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

retical underpinnings of translation and guides practitioners in navigating the intricacies inherent in translating texts across diverse languages and cultures.

Turning to the works in translational linguistics (Baker, 2006; Catford, 1965; Halliday, 2001), its primary goal, unlike comparative linguistics, is to establish *equivalent relationships* between two texts of source language (SL) and the target language (TL). It is noted that the *concept of equivalence* is not central to comparative linguistics, which aims to identify systemic similarities and differences between languages compared. However, translation-based comparison can also be conducted within comparative linguistics (see Jager, 1973:160; Helbig, 1973:173; 1981:82; Muhlner, Sommerfeldt, 1981:79; Schweitzer, 1987:159). These and other issues are detailed in the classic work by Gert Jager, "Translation and Translational Linguistics" (1975), which includes sections devoted to the social role of translation, its history, various types of translation and linguistic mediation, as well as a number of other translation theory issues.

In trying to verify equivalent relationships, the degree to which the original text (source text) and the translated text (target text) correspond to each other in terms of meaning, function, and impact on the reader, researchers use various methods to analyze and compare large bodies of texts (corpora) in both the original and translated languages. The **cluster method** observed in corpus linguistics involves identifying and analyzing patterns of language use (such as collocations, idiomatic expressions, and syntactic structures) within these corpora. By examining these patterns, researchers can assess how closely translations adhere to the original texts in terms of these various dimensions of equivalence, thereby evaluating the effectiveness and accuracy of the translation. This method is instrumental in hypothesis generation by identifying structures within large or complex sets of data, which are otherwise difficult to interpret through direct inspection. Hermann Moisl delves into this methodology. He explains how the cluster method is employed in linguistics for hypothesis generation, involving the abstraction of data from text, data analysis, and the formulation of a hypothesis based on the inference from the results. This approach contributes significantly to a quantitative empirical methodology within the discipline, especially given the challenges posed by the size and complexity of digital corpora in modern linguistics (Hermann, 2015).

The purpose of the article is to present the methodological algorithm of the corpus-cluster method by verifying legal translatemes in English, German, and Polish Corpora (based on JRC-Acquis Sub-Corpus OPUS). This verification process serves several critical purposes in the context of legal translation studies: (a) identification of translatemes; (b) consistency and standardization; (c) comparative legal linguistics; (d) improving translation quality; (e) developing translation tools and resources; (f) enhancing legal communication. Such verification of legal translatemes using the corpus-cluster method in a multilingual

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

context aims to advance the precision, reliability, and efficiency of legal translations, thereby supporting better legal understanding and cooperation across English, German, and Polish-speaking regions.

Background and literature review

The literature on parallel corpora, as depicted in some sources (Hernán Robledo & Rogelio Nazar, 2023) covers a wide range of topics, from the creation of universal corpora to the application of parallel corpora in various linguistic and computational fields. The focus ranges from building extensive language resources to address linguistic diversity (Abney / Bird, 2010, 2011), exploring the Unicode standard for text representation (Allen et al. 2012), to the use of corpora in digital humanities and linguistic analysis (Dipper / Schultz-Balluff 2013; Christodoulopoulos 2013). Scholars like Čermák / Rosen (2012) and Steinberger et al. (2013) discuss the usage of parallel corpora in NLP (Natural Language Processing) applications, including sentiment analysis. Sentiment analysis, also known as opinion mining, is the process of determining the emotional tone behind a body of text. This is a common task in NLP that involves identifying and categorizing opinions expressed in a piece of text, especially to determine whether the writer's attitude towards a particular topic, product, etc., is positive, negative, or neutral.

The historical context and technological advancements in the compilation of linguistic databases are also emphasized (Erjavec 2004; MacWhinney 2000). Overall, these works collectively highlight the evolving nature, utility, and challenges in the field of corpus linguistics and its applications in modern language studies and computational linguistics.

The corpus approach in linguistic research: a critical analysis

In the academic field of linguistic studies, the corpus approach has elicited extensive debate and diverse perspectives. An illustrative example of this discourse is the criticism levelled by Noam Chomsky, a prominent figure in American generative linguistics. In a 2004 interview, Chomsky dismissively characterized corpus linguistics as insignificant, a viewpoint echoed by his contemporaries and followers, such as Professor Lees, who, in 1962 at Brown University, criticised the creation of a corpus as a squandering of governmental resources (Andor 2004). Chomsky's critique primarily hinges on the perception that corpus linguistics is reductive, focusing merely on data observation without

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

substantially contributing to scientific knowledge, particularly in cognitive and practical domains.

Contrasting Chomsky's stance, mid-20th-century linguistic research experienced a shift toward text-centric and discourse-centric approaches, predicated on the notion that language system fragments should be examined through representative text corpora. This shift, however, sparked debates over the criteria for representativeness, a topic still under development.

The necessity and evolution of the corpus approach in applied linguistic research raise critical questions. Preliminary answers suggest that omitting corpus linguistics from modern linguistic science could result in significant regression, a viewpoint shared by scholars like Plungyan (2008, 2009).

The corpus's role in linguistic studies is dissected through three distinct approaches outlined by Bober (2018): the radical-categorical approach, which concentrates on fundamental aspects; the moderate-sceptical approach, offering a balanced perspective; and the scientific-promising approach, underscoring the potential of a corpus in advancing linguistic research. These approaches collectively embody the global scientific community's diverse attitudes towards corpus linguistics as a tool.

Boriskina (2015) presents *radical scholars' arguments*, including Newmeyer (2003), Prodromou (1997), and Chomsky, who dismiss the corpus approach's potential, criticizing its lack of a dedicated Universal Decimal Classification (UDC) index. They liken the corpus to a mere "catalogue," arguing that compiling material as a card index is a standard linguistic practice and does not merit special research status. Boriskina challenges this view, asserting that a marked and annotated corpus offers more in terms of scale, functionality, and research opportunities than a simple catalogue.

The historical context of the corpus approach is vital to comprehend its development. The Oxford English Dictionary, compiled in the late 19th century, utilized numerous authentic example cards collected by lexicographers. Centuries prior, high-status text collections were compiled in libraries for studying rhetoric, style, and grammar. In the 13th century, 500 monks created the first concordance to the Latin translation of the Bible, a precursor to modern corpus methods. The Danish linguist Johansson, representing the Copenhagen School of Structuralism, compiled a vast collection of linguistic examples, indicating a potential inclination towards corpus methods had computational technology been available.

The moderate-skeptical perspective, as represented by scholars like Apresyan, acknowledges the value of corpus technologies but raises concern over potential misrepresentations and misuse of quantitative data. Apresyan argues that merely calculating word frequency is insufficient for definitive linguistic object analysis (Boryskina 2015: 25). This scepticism is compounded by perceived dilettantism

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

among corpus linguists due to the evolving terminological apparatus and methodology of corpus research.

In contrast, *the scientific-perspective approach* in corpus linguistics offers a counter-narrative to radical and moderate skepticism. Prominent in international scientific conferences such as *Corpus Linguistics* and *Dialogue*, this viewpoint focuses on enhancing search engines and information research methodologies within corpora.

Plungyan (2009) asserts that contemporary linguistic research is inseparable from corpus linguistics, an assertion echoed by linguists like J. Sinclair (2006). Meyzerska, in her article *Corpus approach in modern linguistics: perspectives and ways of applying*, underscores the corpus approach's objectivity, verifiability, and efficiency in studying language variants, dialects, and styles, as well as in facilitating historical comparisons, quoting Svartvik (1992: 7).

The corpus is recognized as both a theoretical resource and a practical tool, instrumental in fields like machine translation, speech recognition, and synthesis, and in developing language-related software programs. The scientific-perspective approach advocates for the effective study and interpretation of both frequently and infrequently used language units using corpus materials. Through comparing and analyzing data from various corpora, linguistic variability and patterns can be observed, aiding in predicting future linguistic developments.

Plungyan (2008) highlights the empirical relevance of the corpus-based approach, positing that corpus research innovations pave the way for new *corpus dictionaries* and *corpus grammars*, compiled and verified against a specific corpus. This approach enhances the reliability and verifiability of dictionaries and grammars, reducing subjectivity and incompleteness. The creation of analyzers and specialized dictionaries for automated corpus markup, including morphological, syntactic, and thematic aspects, represents a unique technological feasibility in corpus linguistics (cited in Boryskina 2015: 27).

Cluster method in corpus linguistics with reference to translation studies

Within the scope of corpus linguistics, cluster analysis continues to enjoy widespread popularity. This statistical method, integral for data processing, organizes items into groups, or clusters, based on the degree of their association. Divjak and Fieller (2019) underscore the necessity for analysts to make informed decisions regarding dissimilarity measures and grouping algorithms. These decisions require not only a comprehensive understanding of the data but also an awareness of the theoretical implications involved.

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

Contrasting with the perspectives of Baayen (2008), Johnson (2008), and Gries (2009), the primary objective of cluster analysis is to provide researchers with at least a foundational comprehension of the underlying processes when a dataset is analyzed using a specific cluster analytic technique.

A predominant application of cluster analysis is in classification, where subjects are segregated into groups such that each subject shares more similarities with others in its group than with subjects outside the group. The effectiveness of clustering is measured by two key metrics: (a) *intracluster distance*, indicating the proximity between data points within a cluster, which should be minimal for a strong clustering effect, denoting homogeneity; and (b) *intercluster distance*, the separation between data points in different clusters, which should be substantial in cases of strong clustering, signifying heterogeneity (Qualtrics 2021).

The choice of the cluster analysis algorithm is crucial, especially when dealing with mixed data types. Major statistical packages offer an array of preset algorithms, each tailored for specific data analyses. Mirkes (2011) identifies two algorithms particularly suited for cluster analysis:

K-Means Algorithm determines the existence of clusters by identifying their centroid points. A centroid is the mean of all points in a cluster. The algorithm iteratively assigns each data point to a cluster by evaluating the Euclidean distance to the centroid, which is initially random and evolves with each iteration. K-means is prevalent in cluster analysis; however, its utility is primarily confined to scalar data (Mirkes 2011).

K-Medoids Algorithm. Functionally similar to K-means, K-medoids, instead of mean centroids, use medoids-actual data points from the dataset, making it more interpretable. This approach is advantageous for survey data analysis, accommodating both categorical and scalar data. Unlike K-means, K-medoids can measure distances in multiple dimensions, representing diverse categories or variables (Mirkes 2011).

When managing datasets with a large number of variables, such as complex surveys, simplifying data prior to performing cluster analysis can be beneficial. This simplification, often achieved through factor analysis, reduces the number of dimensions, potentially resulting in clusters that more accurately reflect the true patterns in the data. Factor analysis combines variables that are related to the same underlying concept, thus condensing the data into fewer dimensions. This technique is instrumental in distilling large sets of variables into manageable, interpretable factors.

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

OPUS as an extensive collection of parallel corpora

In the scholarly domain of computational linguistics, the OPUS corpus is distinguished by its extensive scope and diversity. This expansive collection of parallel corpora encompasses over 90 languages, featuring more than 40 billion tokens across 3,800 language pairs, spanning various domains. Remarkably, the corpus contains the largest Spanish-English pair, comprising approximately 36 million parallel sentences. The corpus's augmentation with databases such as ECB, MBS, and OpenSubtitles2011, particularly the latter, is notable for its substantial size and rich linguistic diversity. OPUS's role extends beyond mere data provision; it offers critical tools for processing both parallel and monolingual data, thereby contributing significantly to advancements in machine translation and linguistic research.

The utility of the OPUS corpus in computational linguistics is further evidenced by its application in projects like OPUS-MT. As detailed in *OPUS-MT – Building open translation services for the World* by Jorg Tiedemann and Santhosh Thottingal (2020), the corpus plays a crucial role in facilitating machine translation across a broad spectrum of languages, including those with limited digital representation. Its impact is not confined to the development of neural machine translation models, as elucidated in the 2020 publication. The corpus serves as an indispensable resource in linguistic research and education, offering unparalleled access to large-scale linguistic datasets that are instrumental in advancing language studies. Furthermore, OPUS's contribution to the field of artificial intelligence, specifically in the realm of machine translation, is underscored by its support for open-source initiatives, highlighting the democratization of language technology.

In the landscape of corpus translation studies, parallel corpora emerge as invaluable resources for linguistic research and natural language processing (NLP) applications. A primary application of these corpora is in statistical machine translation (SMT), where extensive aligned data are standardly employed to learn word alignment models between the lexica of two languages, exemplified in the Giza++ system by Och and Ney (2003). Moreover, parallel corpora are instrumental in the projected learning of linguistic structure in NLP. This methodology utilizes supervised data from a resource-rich language to guide unsupervised learning algorithms in a target language. While some techniques may not necessitate parallel texts (e. g., Cohen et al. 2011), the most efficacious models typically employ sentence-aligned corpora (Yarowsky / Ngai 2001).

Christos Christodouloupoulos (2013) emphasize the need for extensively translated texts to access parallel material in diverse languages. Following the methodology of Resnik et al. (1999), they have created a massively parallel corpus based on Bible translations (cf. Abney / Bird 2010, 2011). The United Bible

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

Societies (2013) report indicates that the Bible, with at least 2,527 translations of parts and 475 complete translations, surpasses all other literature in translation volume, a testament to its global reach and linguistic diversity.

The Acquis Communautaire (AC), encompassing the entirety of European Union (EU) law applicable in EU Member States, serves as another critical source of parallel texts. This evolving legislative collection, spanning from the 1950s to the present, is available in 24 languages of the EU's 27 Member States, excluding Irish. As Steinberger et al. (2006, 2014) note, the Acquis Communautaire is not only the largest parallel corpus in existence when considering both its size and the extensive number of languages involved, but it also stands out for its unique range of rare language pair combinations, such as Maltese-Estonian and Slovenian-Finnish.

The value of parallel corpora in cross-lingual research is immeasurable, growing with their size and the diversity of languages they encompass. While some language pairs are well-represented in parallel corpora, many others remain underrepresented, highlighting the importance of resources like the Acquis Communautaire in bridging these linguistic gaps. The corpus's accessibility and the rarity of language pairings it offers make it an exceptional asset in the field of computational linguistics.

Legal translateme as the key unit in the theory and practice of legal text translation

The concept of the *translateme*, as expounded in Tulyenev's *Theory of Translation* (2004: 89–103), plays a pivotal role in the translation process, especially in fields requiring high precision and contextual sensitivity, such as legal translation. It means that a translateme is defined as the fundamental unit of translation, encapsulating a blend of content and its linguistic expression as it moves from the source language (SL) to the target language (TL). A translateme is not merely a linguistic construct; it represents a synergy of the content expressed in the SL and the same content articulated in the TL, thereby enabling translators to capture not just the literal content but also the underlying contextual and cultural significance of the original text.

The size and scope of a translateme can vary greatly, extending from minute elements like sounds or letters to more substantial segments, including phrases or even entire texts. This variability is largely contingent on the typological relatedness between the SL and TL and the intrinsic characteristics of the text, particularly in genres that are heavily nuanced, such as poetry or artistic works. The translation process involves the initial identification of translatemes, which

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

are then potentially modified or refined within the broader syntactic structure of the target language. This meticulous verification process at the denotative level is crucial to ensure that the translation not only mirrors the literal meaning but also encapsulates the deeper, more nuanced contextual significance, thus achieving both adequacy and relevance for the target audience.

In the specific context of legal translation, the translateme assumes even greater significance. Here, it represents the smallest unit of translation, embodying both the content and its linguistic expression while transitioning from the SL to the TL. Legal translatemes are particularly complex due to their need to encompass not only the literal meanings but also the deeper contextual and cultural nuances that are integral to legal texts. Their complexity is further heightened by their variability, which can range from single terms to intricate phrases or clauses. This variability is influenced by the typological similarities or differences between the SL and TL, as well as the specific legal terminologies and concepts prevalent in the texts. The translation process in this context involves an initial identification of legal translatemes, followed by necessary modifications or restructuring to align with the sophisticated syntactic and semantic structures of legal language. The verification of these translatemes at a denotative level is imperative to ensure that the translated text accurately reflects the legal realities and maintains fidelity to both the spirit and letter of the original legal document. This rigorous translation process is essential to ensure that the final translated legal document is not only linguistically accurate but also legally operative and contextually relevant in the TL.

The concept of the translateme, as delineated in Tulyenev's seminal work, is integral to the translation process, particularly in fields like legal translation, where precision, contextual understanding, and cultural sensitivity are paramount.

Methodology of corpus verification of legal translatemes in the source and target languages using OPUS, JRC-Acquis Sub-Corpus

The methodological approach to employing the *cluster-corpus translation method*, particularly through the use of the JRC-Acquis sub-corpus of the OPUS corpus, offers a structured and precise framework for legal translation. This corpus, encompassing a wide range of legislative texts from the European Union dating from the 1950s to the present, provides a rich resource for legal translators. The cluster method in this context involves *classifying legal translatemes* and analyzing them in terms of their *structure*, *meaning*, and *key translation methods*. The process can be elucidated as follows:

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

I. Selection and classification of legal L1 translatemes

The initial stage involves selecting legal translatemes from a source language (L1), such as English. These translatemes are then classified based on their components, which may be single, dual, or triple elements. The selection process can be facilitated through the Advanced Query function in the JRC-Acquis sub-corpus. For instance, by entering the adjective *bilateral* in the search field, a range of relevant contexts and translatemes can be retrieved.

II. Identification of multi-component legal translatemes

The next step is to sift through the search results to identify multi-component legal translatemes that incorporate the adjective *bilateral* and correspond to both L1 and the target languages (L2), such as German and Polish.

This process helps in pinpointing the most commonly used legal translatemes in L1 that feature the *bilateral* cluster along with right-sided valence components.

III. Analysis of translateme components

Once the relevant translatemes are selected, they are analyzed to understand their structural and semantic compositions. This analysis involves examining the various components of the translatemes, such as one-component translatemes in the context of their use in legal texts.

The focus on one-component translatemes can reveal insights into how a specific term, like *bilateral*, is employed within the legal framework of the EU texts and how it translates across different languages.

IV. Verification against real-world context

Crucial to this method is the verification of the translatemes against the real-world context or denotatum. This step ensures that the translation is not only linguistically accurate but also contextually appropriate and aligned with the legal realities represented in the text.

The comparison of the text with the corresponding denotat is essential to avoid common errors, particularly those made by novice translators who may interpret the text in isolation from its real-world application.

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

V. Ensuring legal nuances and contextual appropriateness

By utilizing resources like the JRC-Acquis sub-corpus, translators can validate the accuracy and contextual appropriateness of their translations. This validation is crucial for preserving the legal nuances of the original text in the translated version.

The application of the *cluster-corpus translation method* using the JRC-Acquis sub-corpus provides a systematic approach for the classification and analysis of legal translatemes. This method emphasizes the importance of context in legal translation, ensuring that translations are not only linguistically accurate but also legally and contextually relevant.

In Table 1, *bilateral cluster* is presented, which showcases the nuanced application of this specific term within various legal contexts. This example highlights how a singular term like *bilateral* carries multifaceted implications in legal language, demonstrating the intricacies involved in accurately translating and interpreting such terms across different legal systems and languages. The table may further illustrate the variations in meaning and usage of *bilateral* in diverse legal documents, reflecting the term's adaptability and significance in international legal discourse.

Tab. 1: One component legal translateme

EN	… and certification conditions are laid down in **bilateral** agreements between the Union and third countries
DE	Wenn in **bilateralen** Abkommen zwischen der Union und Drittländern besondere Bedingungen hinsichtlich der Tiergesundheit und der Bescheinigungen festgelegt sind, so gelten diese an Stelle der Bedingungen in Absatz 1.
PL	W przypadku gdy w umowach **dwustronnych** zawartych między Unią a państwami trzecimi określono specjalne warunki dotyczące zdrowia zwierząt i warunki dotyczące świadectw, stosuje się te warunki w miejsce warunków określonych w ust 1.

Source: compiled by the authors

In the analysis of the translations of the phrase *health and certification conditions are laid down in bilateral agreements between the Union and third countries*, particular attention is paid to the use of the word *bilateral* in both German and Polish translations. This analysis will explore how the word *bilateral* is incorporated into the respective languages, considering the structural and semantic aspects of the translations, as well as any notable transformations.

The phrase in English sets a context where specific conditions regarding animal health and certification are established in agreements that are bilateral – meaning they involve two sides, in this case, the Union and third countries.

The German translation is: *Wenn in bilateralen Abkommen zwischen der Union und Drittländern besondere Bedingungen hinsichtlich der Tiergesundheit*

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

und der Bescheinigungen festgelegt sind, so gelten diese an Stelle der Bedingungen in Absatz 1. In this translation, *bilateral* is rendered as *bilateralen*, which is the inflected form of *bilateral* in German, used to agree with the noun *Abkommen* (agreements). The structure of this sentence in German closely follows the English structure, with the word *bilateral* directly modifying *Abkommen.* This reflects a straightforward and direct translation approach, where the grammatical and lexical properties of the word *bilateral* are preserved, ensuring that the meaning of mutual agreement between two parties (the Union and third countries) is accurately conveyed.

The Polish version, *W przypadku gdy w umowach dwustronnych zawartych między Unią a państwami trzecimi określono specjalne warunki dotyczące zdrowia zwierząt i warunki dotyczące świadectw, stosuje się te warunki w miejsce warunków określonych w ust 1*, translates *bilateral* as *dwustronnych*. This word is a form of *dwustronny*, which literally means *two-sided* and is the equivalent of *bilateral* in Polish. The placement of *dwustronnych* before *umowach* (agreements) is a slight structural shift from the English original but is in line with Polish grammatical structure. This translation captures the essence of the agreements being between two parties, retaining the semantic integrity of the original phrase.

In both German and Polish translations, the word *bilateral* is effectively translated to convey the notion of agreements involving two entities – the Union and third countries. The German translation *bilateralen* and the Polish *dwustronnych* both serve to accurately represent the concept of bilateral agreements. Structurally, the German translation closely follows the English syntax, while the Polish translation adapts to the syntactic norms of the Polish language. In terms of meaning, both translations maintain the original concept of mutual agreements, demonstrating the importance of accurate translation in preserving the essence of legal and technical terminology across languages.

In Table 2, *bilateral agreement* is presented, which demonstrates the intricate linguistic and legal interplay between the source and target languages in the process of translation, highlighting the nuances involved in accurately conveying the term across different legal systems.

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547

Tab. 2: Two-component legal translateme

EN	of Cape Verde should conclude, without delay, a **bilateral agreement** on facilitating the issue of short-stay
DE	Daher sollten das Königreich Dänemark und die Republik Kap Verde unverzüglich ein **bilaterales Abkommen** zur Erleichterung der Erteilung von Visa für einen kurzfristigen Aufenthalt mit ähnlichen Bestimmungen schließen, wie sie das Abkommen zwischen der Union und Kap Verde enthält.
PL	W związku z tym wskazane jest, aby władze Królestwa Danii i Republiki Zielonego Przylądka zawarły niezwłocznie **dwustronną umowę** o ułatwieniach w wydawaniu wiz krótkoterminowych na warunkach podobnych do zawartych w umowie między Unią a Republiką Zielonego Przylądka.

Source: compiled by the authors

In this analysis, we will examine the translations of the phrase involving *a bilateral agreement* into German and Polish, with a focus on how the term *bilateral agreement* is translated and integrated. We will consider the structural and semantic aspects, and any translation transformations that occur.

The phrase in English suggests that the Kingdom of Denmark and the Republic of Cape Verde should promptly establish a bilateral agreement to facilitate the issuance of short-stay visas. The term *bilateral agreement* here is pivotal, implying an arrangement involving two parties.

The German translation reads, *Daher sollten das Königreich Dänemark und die Republik Kap Verde unverzüglich ein bilaterales Abkommen zur Erleichterung der Erteilung von Visa für einen kurzfristigen Aufenthalt mit ähnlichen Bestimmungen schließen, wie sie das Abkommen zwischen der Union und Kap Verde enthält.* Here, *bilateral agreement* is translated as *ein bilaterales Abkommen*. The structure of this sentence in German closely follows the English structure, with *bilaterales Abkommen* directly corresponding to *bilateral agreement*. This direct translation approach maintains both the grammatical and lexical properties of the English phrase. The term *bilaterales* (bilateral) aptly conveys the two-party nature of the agreement, ensuring the original meaning is accurately reflected in the translation.

In Polish, the phrase is translated as *W związku z tym wskazane jest, aby władze Królestwa Danii i Republiki Zielonego Przylądka zawarły niezwłocznie dwustronną umowę o ułatwieniach w wydawaniu wiz krótkoterminowych na warunkach podobnych do zawartych w umowie między Unią a Republiką Zielonego Przylądka.* The term *bilateral agreement* is translated as *dwustronną umowę*, where *dwustronna* is the Polish equivalent of *bilateral*, and *umowę* means *agreement*. This translation captures the essence of an agreement involving two parties. The structural placement of *dwustronną umowę* in the Polish sentence aligns well with the syntactic norms of the Polish language, and despite

ISBN Print: 9783847117544 – ISBN E-Book: 9783847017547