



## METHODOLOGY OF PHILOLOGICAL STUDY: THE NEW DIRECTION IN APPLIED LINGUISTICS

*Assoc. Prof. Yulia DEMYANCHUK*

*Lviv State University of Life Safety, Ukraine*

### **Abstract**

The central novel discovering of this paper is that we propose a new direction in philology – corpus-applied translation studies. In contrast to the traditional philology corpus-applied translation studies is the sphere of historical, literary, texts of official and business documents, and other texts involve a specific application methodology. This methodology comprises procedures aimed at decoding the meaning of the analyzed phenomenon, narrowing down the range of possible meanings. The initial stage of applying corpus-applied methodology involves identifying dominants from the corpus continuum of various genres, followed by interpreting their content embodied in a system of meanings that are semantically significant. Under these conditions, utilizing corpus data and analyzing their structure and context contribute to enhancing translation quality, ensuring precise reproduction of content, style and specificity across various genres of texts. Reorganizing corpus-applied aspects through the extraction of dominants from historical, literary, or official-business genres shifts the focus on the systematic trends, favoring linguistic means of this subcategory oriented towards establishing a hierarchical structure of content and the scope of concepts. An integrated approach to analyzing texts of a specific style through the prism of corpus-applied research principles involves considering various factors that elucidate its semasiological perspective and influence its implementation in the text. The combination of corpus-applied methods and analytical techniques aids in understanding multi-genre texts as verbal reflections of language in all its manifestations.

**Key words:** linguistic methodologies, corpus-applied translation studies, traditional philology, corpus analysis, base language, target language

### **Introduction**

In the 21<sup>st</sup> century, the field of philology stands at a crossroads, balanced between its historical origins and the ever-evolving view on modern linguistic challenges (Naudé 2017). Philology, originating from the Greek words *philos* (love) and *lógos* (word), traditionally covered the study of language, literature and culture through historical texts (Liddell and Scott 2017). However, the complexities of



contemporary society demand an adaptive approach to meet the multifaceted challenges facing language studies. Philology, in its traditional sense, referred to the comprehensive study of language, literature, history and culture through the examination of different texts (Crystal 2023). Philologists meticulously analyzed language structures, traced etymologies and deciphered the historical and cultural significance of texts to gain insights into past civilizations (Errington 2008). As stated by J. Errington, traditional linguistic methodologies within philology involved detailed examination of written texts, often in ancient languages (Ibid). Comparative analysis, textual criticism formed basis of philological studies. Scholars relied on close reading, decipherment of scripts and the application of grammatical and linguistic frameworks to interpret and contextualize the diversity of documents. Nevertheless, the 21<sup>st</sup> century presents a dynamic set of challenges that require an evolution in linguistic methodologies and extended linguistic directions. Rapid advancements in technology, globalization, new cultural paradigms and the digital revolution have reshaped the scientific language environment. The emergence of new linguistic forms, the blending of languages have created the modern platform that necessitates novel approaches in linguistic studies.

The modernization within philology is also needed to navigate the contemporary linguistic complexities. This discipline must evolve to address not only historical texts but also new language variations, digital communication, socio-cultural influences and the impact of technology on language evolution. The study of philological processes in the 21<sup>st</sup> century extends beyond the boundaries of deciphering ancient manuscripts. It encompasses the exploration of digital footprints, the analysis of online discourse and the preservation of endangered languages in a rapidly changing world. Fundamentally, this field requires innovative, dynamic approach to capture the diverse sides of language in the contemporary era. Thus, the adaptation of traditional philology becomes imperative (Cenk 2022, 31–49). This adaptation signifies not the refuse of historical methods but rather their augmentation with modern tools, technologies in the context of the interdisciplinary perspectives.

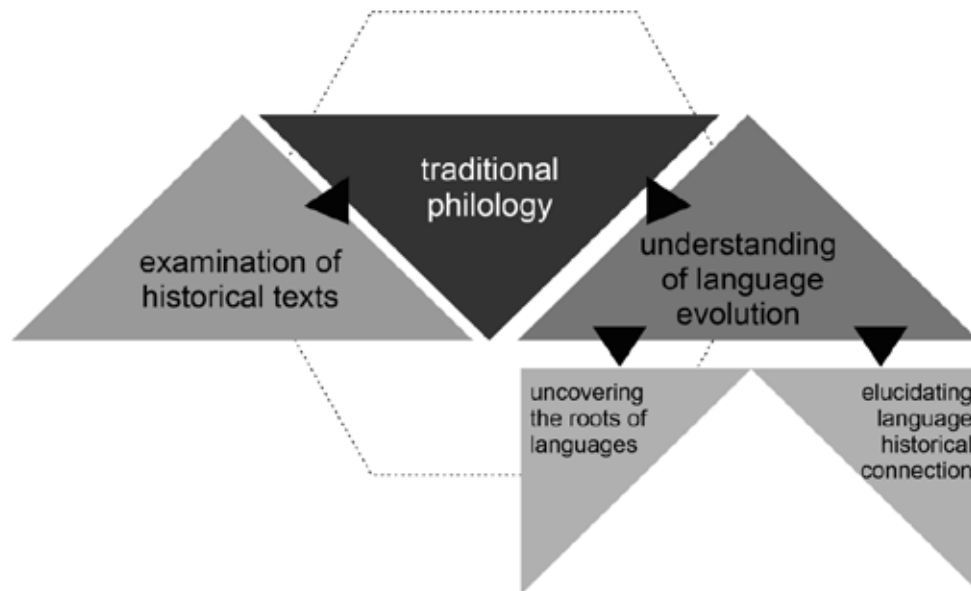
This paper aimed to reveal the main features of linguistic evolution, cultural context, linguistic methodologies and to propose the new direction *corpus-applied translation studies* within philology.

## **Results**

### **1. The role of traditional philology in language studies**

The evolution of linguistic studies has been steeped in the traditions of rigorous analysis, comparative studies and textual interpretation. In accordance with J. M. Power, traditional linguistic methodologies, foundational to the development of philology and are focused on precise examination of historical texts (Power 2020). These methodologies have laid the framework for the comprehensive understanding of language evolution and cultural heritage. Traditional philology played a key role in uncovering the roots of languages and elucidating their historical connections (Ibid). Researchers were engaged in detailed analysis of ancient manuscripts, employing methods such as textual criticism, etymological analysis and comparative linguistics. Through these methods, philologists reconstructed

linguistic lineages and traced the historical aspects of languages, contributing significantly to our understanding of linguistic evolution (Jäger 2019, 151–182).



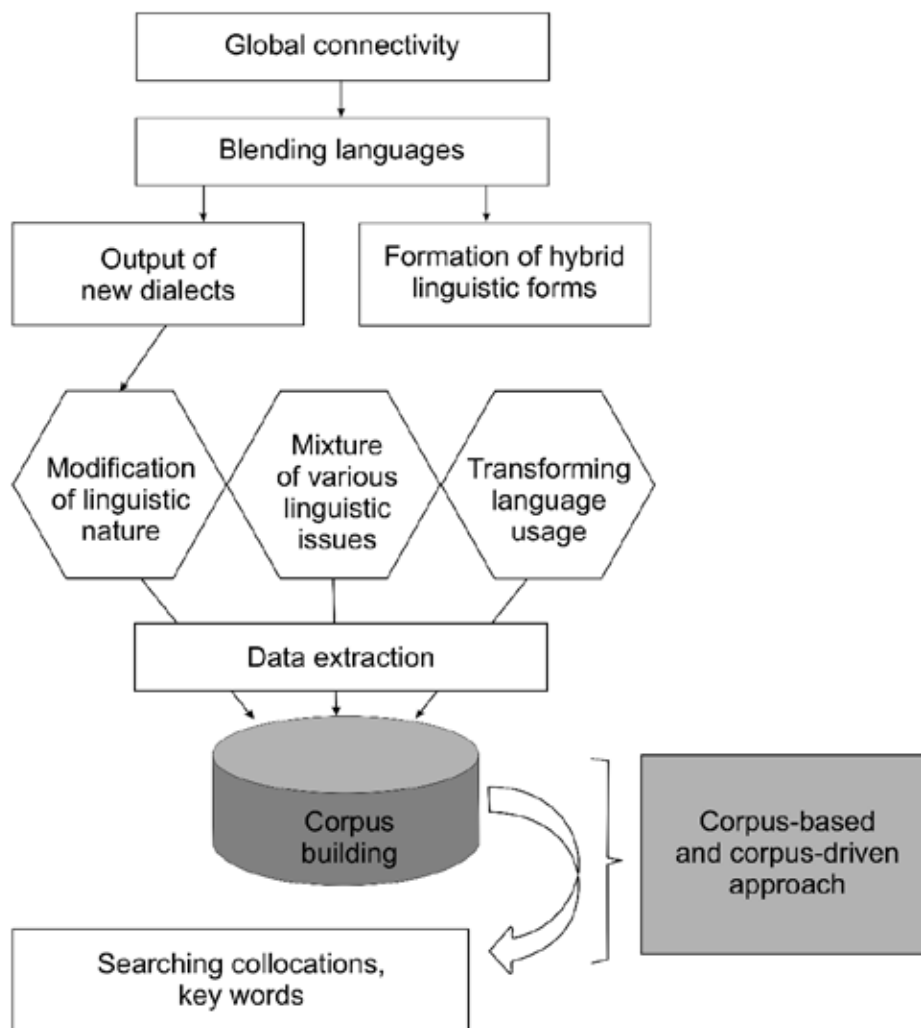
**Fig.1.** Aims of traditional philology

Moreover, traditional philology extended beyond the simple decipherment of texts. It becomes a tool through which to view the influence of past societies, cultures on modern linguistics. As claimed by D. Slade, the study of historical texts provided invaluable insights into the socio-cultural contexts, religious practices, philosophical beliefs of bygone eras (Slade 2020). By examining linguistic aspects embedded within texts, philologists solved not only language phenomena but also the specifics of ancient civilizations. Classical linguistic approaches, including historical linguistics, comparative linguistics and textual analysis made profound contributions to the field. Historical linguistics focused on tracing the evolution of languages over time, identifying linguistic changes and reconstructing proto-languages. Comparative linguistics facilitated the comparison of languages, identifying commonalities and divergences to establish language families and linguistic relationships. Textual analysis is one from the most important linguistic approaches, involved the detailed examination of texts for linguistic patterns, stylistic elements, grammatical structures. This method allowed scholars to reconstruct lost languages and preserve linguistic heritage. Additionally, classical linguistic approaches improved a deeper understanding of language structure and typology. This approach developed the way for modern theories in linguistics. The systematic examination of historical linguistic data provided a strong foundation upon which contemporary linguistic studies have been built.

## **2. Emergence of 21<sup>st</sup> century challenges in linguistics**

The 21<sup>st</sup> century has introduced in a new era marked by a multitude of challenges that significantly impact linguistics. These challenges transcend traditional boundaries and demand an

adaptive approach to comprehend the complexities of language in the modern world (Muñoz-Basols, Neville, Lafford and Godev 2023, 171–194). The coming of technology and globalization has profoundly altered linguistic landscapes (Eliezer and Miriam 2015, 19–37). Communication is no longer limited of geographical boundaries. It overpasses borders through digital platforms and online interactions. Global connectivity has led to the blending of languages, the output of new dialects and the formation of hybrid linguistic forms. Cultural exchanges, migration, intercultural interactions have contributed to the modification of linguistic nature, reflecting the mixture of various linguistic issues (Geraghty and Conacher 2014). Moreover, the rise of digital media, social networking and the internet have transformed language usage and communication patterns. Abbreviations, emojis, internet slang have become integral components of contemporary communication, changing conventional linguistic trends. This digital evolution poses unique challenges to linguists in understanding and interpreting these linguistic phenomena (see Fig.2).





**Fig.2.** Challenges in linguistics

According to the researchers L. Amelia, Lolitha and N. Balqis, digitalization has fundamentally reshaped language and communication patterns (Amelia and Balqis 2023, 544–556). Online platforms, social media, digital communication tools have introduced a new lexicon and communication style (Ibid). The spread of digital footprints and user-generated content influences online discourse and linguistic analysis. It presents numerous platforms for linguistic studies.

### **3. The new direction in philology: corpus-applied translation studies**

Contemporary philologists navigate the complexities of modern linguistic challenges by employing interdisciplinary approaches. They integrate insights from sociolinguistics, computational linguistics and cognitive science, allowing for a more comprehensive understanding of language usage. This adaptation involves information technology as a tool for linguistic analysis, using computational methods for text mining and exploring linguistic corpora to uncover patterns within digital communication (Bednarek and Carr 2020).

The evolution of linguistic methodologies manifests in various innovative approaches that enrich linguistic studies. Computational linguistics, for instance, utilizes computational tools and algorithms to process and analyze extensive linguistic datasets. This field has revolutionized natural language processing, enabling applications such as machine translation, sentiment analysis, chatbots. Furthermore, socio-cognitive approaches explore the intersection of language and cognition, investigating the way social factors influence language use and comprehension. This interdisciplinary field bridges the gap between psychology and linguistics, shedding light on the cognitive mechanisms behind language acquisition. Another perspective area is corpus linguistics, which operates large collections of linguistic data to examine language patterns, usage variations, semantic structures. This field allows linguists to captivate empirical evidence from extensive datasets, providing a deeper understanding of language in different contexts. Additionally, the integration of digital humanities and linguistics has facilitated the analysis of digital archives, social media, and online communication patterns. Linguists use data-driven methodologies to decode linguistic phenomena in digital spaces, revealing socio-cultural variations within online interactions.

The formation of corpus linguistics as a research direction is associated with English-language linguo-culture of the 20<sup>th</sup> century. The later theorization of this direction is explained by the insufficient conceptual apparatus due to differences in the interpretation of basic branch beliefs in various discursive practices (Demyanchuk 2023, 53–58). The concept of corpus linguistics has been developed to denote a division of linguistics, the main tasks of which include the investigation, elaboration, creation and use of textual corpora, as well as linguistic analysis based on them (Fellbaum 2019, 735–767). However, the key area of corpus linguistics research is to define the general principles of constructing, processing, and using data from linguistic corpora (Stefanowitsch 2020).



The fundamental principles of defining a text corpus involve the identification of various concepts within modern Corpus Linguistics (Teubert and Cermakova 2007, 153), shaping both theoretical strategies and practical principles for its utilization (Bennett 2010). In general linguistics, a text corpus refers to a collection of linguistic data from written texts, which can serve as a starting point for linguistic description or as a means of verifying language hypotheses (Crosthwaite, Ningrum and Schweinberger 2022). In the specific context of Corpus Linguistics, a corpus of texts is a compilation of a large variety of written and oral text samples available in machine-readable form, systematically collected for the presentation and utilization of language diversity (Stefanowitsch 2020). Many scholars refer to a text corpus as a collection of texts or their fragments gathered for linguistic research (LR) purposes (Morgan 2022). The theory covers both pre-electronic and electronic types of linguistic corpora (LC) (Hasselgren 2020, 143–174). Pre-electronic corpora refer to those created before the computer era that served as the basis for manual analysis of linguistic material: the determinative feature of a pre-electronic corpus is not its machine-readable form but its direct purpose, involving a text-oriented approach to language study [Ibid].

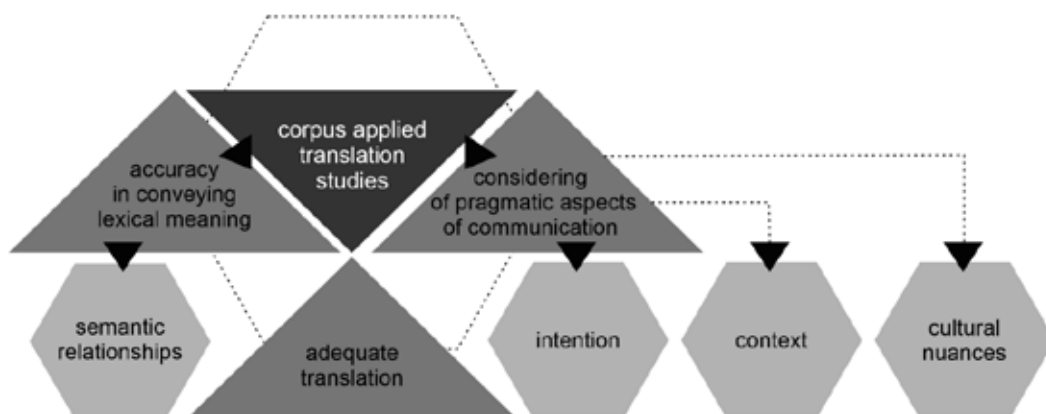
Modern corpus analysis researchers adhere to a more specific definition, as described by D. Stifter, F. Qiu and M. Aquino-Lopez (Stifter, Qiu and Aquino-Lopez 2022, 529–553): they interpret a corpus of texts as a collection of authentic texts or their fragments intended for linguistic processing, generally focusing on natural language data rather than artificially created or discovered linguistic material (Dash 2005, 445). Accordingly, text corpora constructed by selecting text fragments to study specific linguistic phenomena contribute to preserving their authenticity in the truest sense (Meyer 2008). At this stage, a text corpus can be understood as the entirety of all texts in a specific language available to a translator (Bisazza and Tump 2018, 2871–2876). From this perspective, the concept of a text corpus entails significant parameters for automated translation: 1) electronic format texts are deemed relevant; 2) the delineation can encompass both linguistic information and the achievement of intended goals; 3) successful analysis enables material verification, where the constructional tool involves the genre, theme and specialization of the text.

The investigation of the translation process in corpus linguistics necessitates considering stylistic differences between languages and finding correspondences between the styles of the original and the translation. Search strategies, methods, techniques and procedures for analyzing source and target texts lean toward translation studies, corpus linguistics, applied linguistics and structural linguistics as components of the corpus-translation studies scientific field. The expansion of translation in various conceptual structures of corpus linguistics entails shaping grammatical norms and translation rules from one language to another, ensuring the grammatical accuracy and comprehensibility of the translation. Taking into account pragmatic aspects helps the translator convey not only the literal meaning of the text but also its implicit meanings, intonation, genre-specific features and other factors that influence the audience's perception and reaction. The penetration of corpus studies into various conceptual structures of applied linguistics leads to the formation of corpus-categorical clusters with

translation components, among which the field of corpus-applied translation studies stands out. **Corpus-applied translation studies** represent a distinct branch of applied linguistics that combines corpus analysis methods and approaches with the study of fundamental translation doctrines. Its focus is on applying corpus data and information technologies in translation studies and translational practice.

The main goal of corpus-applied translation studies lies in understanding translation phenomena based on data collected from various text corpora. Corpus data in the context of translation studies refer to large collections of texts that enable translators and researchers to analyze translations from reports, expert opinions, recommendations, resolutions to international documents, projects, directives, statutes, laws, as well as literary, colloquial, confessional, scientific, journalistic, and official-business texts.

**Fundamental principles of corpus-applied translation studies** encompass: accuracy in conveying lexical meaning alongside the development of synonymy, antonymy and other semantic relationships within the target language text corpus; balanced rendering of thought, idea congruent with literary, colloquial, scientific, journalistic, official and business, confessional styles in the source text corpus; the use of grammatically normative structures in the translation text corpus; focus on phonological aspects of language (sounds, rhythms, intonation, etc.) in both the source and target languages; consideration of pragmatic aspects of communication (intention, context, cultural nuances) within the text corpus; seeking appropriate equivalents (expressions, lexemes, phrases) and reproducing both semantic and stylistic features of lexical elements within the context of corpus-based translation; accurate representation of sentence syntax contributing to an adequate translation; ensuring accuracy and consistency in terminological translation. In corpus-applied translation studies, a comprehensive set of approaches is relevant. In this research we will outline the essence of each one.



**Fig.3.** Taxonomy of fundamental principles of corpus-applied translation studies

Corpus analysis of the base language, as a methodological approach, involves using corpus data in the investigation and analysis of the *base language* (source language). Identifiable criteria for the base language include the translation of primary and intermediate levels of language-lexical,



grammatical, stylistic and phraseological. Corpora encompass textual materials of various genres, styles, themes, communicative situations, selected from diverse sources such as literary texts, periodicals, web pages, colloquial texts. By examining data from the corpus of base language texts, researchers identify typical and unprecedented patterns and tendencies in the usage of words and terminology, develop lexico-grammatical patterns, trace stylistic differences, evolve models of linguistic functioning and devise precise, effective strategies for achieving quality translation. This approach stands as a valuable tool in translation studies by enabling a meticulous analysis of linguistic nuances and usage of linguistic units in the source language. Simultaneously, employing corpus analysis of the base language in applied linguistics and translation studies facilitates not only relative assessments and perceptions of linguistic activity but also provides objective data regarding the development of linguistic competence.

Corpus analysis of the *target language*. The essence of this methodological approach lies in utilizing corpus data as a significant factor in studying and analyzing linguistic peculiarities of the target language, including various linguistic aspects such as aesthetic, psychological, cultural, pragmatic. The multitude of choices in data corpus expands translators' possibilities. As a result of this diversity within corpus-applied translation studies, researchers obtain a vast array of patterns and empirical data for analysis. Corpus analysis of the target language aids translators and researchers in better comprehending and considering linguistic differences between the source and target languages, influencing precise and adequate translation. Additionally, it assists identify common errors or difficulties and subsequently addresses them. The ontology of corpus analysis of the target language is both theoretical and conceptual, aiming to systematize cultural, social, stylistic preferences, and other contextual factors influencing the perception and comprehension of translation.

According to the corpus-applied approach, corpus analysis of the target language is viewed as performing specific tasks related to identifying and analyzing lexical features that differentiate the target language from the source language. Expanding the understanding of corpus analysis of the target language as a corpus-integrated translational construct involves not only presenting speech patterns in the target language but also extracting new expressions, phraseologies, constructions, and other linguistic elements that personalize the paradigm of contemporary corpus-applied translation studies.

**Corpus-applied translation comparison.** This methodological approach is based on using text corpora and applied methods of theoretical, componential, empirical analysis to compare general scientific and specialized, connotative and denotative meanings of lexical units in two or more languages. The gnoseological function of corpus-applied translation comparison is associated with examining various translation variants of parallel texts in the source and target languages, identifying typical errors and seeking optimal translation solutions. A significant feature of corpus-applied translation comparison is its orientation towards specific texts and quantifiable data collected in the text corpus. This approach is also grounded in the interaction of a representative dataset in the realm of corpus translation studies. Given that the concept of corpus-applied translation comparison is a product of



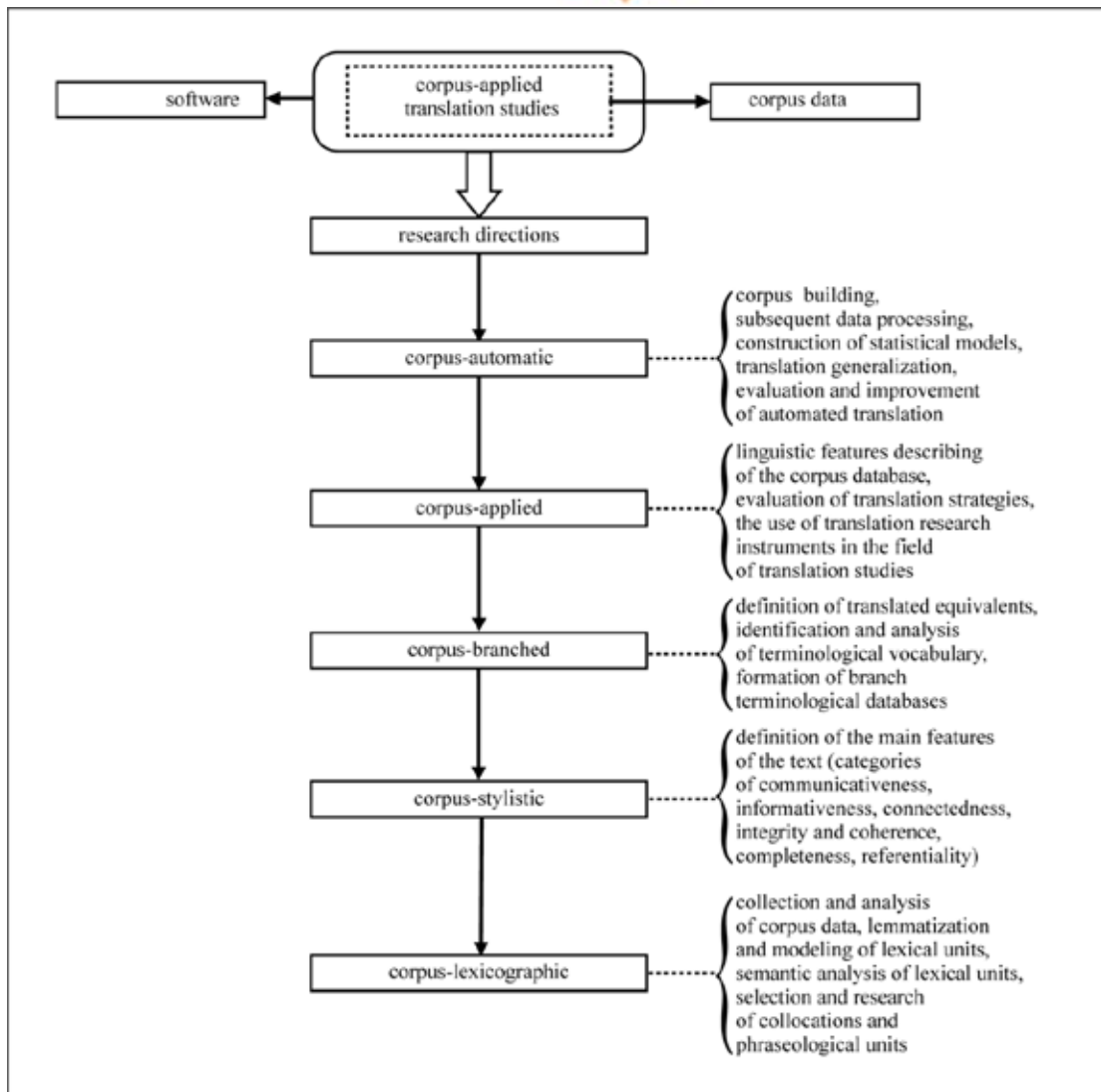


corpus-applied translation studies, reflecting translation issues in applied linguistics combines both linguistic and extralinguistic features, tools and models for extracting and processing data from a text corpus.

Thus, the corpus-applied approach to translation comparison relies on automatic algorithms for finding parallel segments, identifying lexical, grammatical, and stylistic features, analyzing frequency and usage of certain lexical units. The application of such methods ensures linguistically constructed and systematic analysis. One of the key aspects of corpus-applied translation comparison involves engaging translators and linguists in the process of analyzing and interpreting the obtained data. Their knowledge and expert assessment represent a specific type of pragmatic interaction, where the discovery of linguistic phenomena and contextual translation features remains relevant. The combination of representative data from the text corpus and objective expert evaluation contributes to a deeper understanding of the translation process and enhances the quality of translation products.

#### **4. Corpus-applied translation studies: key research directions**

The new direction encompasses various aspects and tasks in the field of translation, utilizing the expressive attribute of a corpus of texts, which, in turn, increases the characterizing and identifying role of the text in corpus-applied translation studies. These study includes corpus-automated, corpus-applied, corpus-branched, corpus-stylistic and corpus-lexicographic directions.



**Fig.4.** Five-level architecture of research directions in corpus-applied translation studies

**Corpus-automated direction:** a prerequisite for this research direction is the development of an automated translation system, typically involving corpus data and machine learning to enhance the quality of translation in terms of accuracy and speed. As rightly observed by researchers R. Aharoni, M. Johnson, O. Firat, modern machine translation technology is grounded in the analysis of a vast volume of texts in both languages represented in corpora (Aharoni, Johnson and Firat 2019, 3874–3884). Accordingly, the scholarly foundation for interpreting text, its linguistic correspondences and translation variants is oriented toward machine learning methods and statistical modeling.

The procedure for applying the corpus-automated principle in corpus-applied translation studies involves several stages. Among them: 1) constructing a corpus of texts, which requires experimental analysis and serves as the basis for creating a translation database; 2) preprocessing the data, involving



various techniques such as tokenization, lemmatization and stop-word removal applied to the classified sample to enhance the quality of translation; 3) building statistical models, one of which is the *n-gram* model, allowing the analysis of word sequences and their frequency in the text corpus to determine the most probable translation options; 4) translation generation, demonstrating the sequence of main statistical models based on algorithms for automated translation; 5) evaluating and improving automated translation implemented through metrics such as translation accuracy, comprehensibility and fluency. Expert translators play a significant role in assessing and correcting shortcomings in automated translation.

The pragmatic implications of the corpus-automated principle in corpus-applied translation studies are outlined by two concepts: *firstly*, the desire to enhance translation models, primarily facilitated by the involvement of additional corpus data, refinement of software algorithms, the use of contextual models, or the application of deep learning; *secondly*, continuous improvement based on feedback from users and translation experts.

**Corpus-applied direction.** Studying the specifics of this direction is a crucial factor in preparing future translators and developing their professional competence in translation activities. This research direction is oriented towards finding optimal means of implementing corpus approaches in solving practical translation tasks (grammatical-translation and textual-translation), driven by the need for rapid, accurate and high-quality translation in the modern world. This can be achieved by using corpus data and applied linguistics tools. The corpus-applied direction also serves as a foundation in software localization, utilizing specialized text corpora for translation and adapting information programs to interlingual semantic-structural correspondences.

The corpus-applied direction generates an increase in scientific studies in various aspects, among which are: 1) the study of linguistic features (stylistic variations, grammatical structures, semantic nuances and lexical resources based on corpus data analysis). For instance, students explore the language implied and encoded within the internal structure of linguistic units in the corpus database, allowing them to better understand the heuristic potential of linguistic representation within the lexical-semantic structure of individual words and expressions; 2) the development of translation strategies, which assists students in developing skills for selecting optimal translation tactics based on the analysis of translated texts in the corpus; students study different translation options, compare their effectiveness, and choose the most valid option for a given context; 3) the utilization of various translation tools based on corpus data to enhance technical methods of text reproduction; this process aims to assimilate automated translation programs, terminological databases, concordances, and other tools that enable finding necessary translation options, identifying semantic complexity, and expressing semantic correspondence in translation; 4) research in the field of translation studies, prompting students to conduct experimental studies using corpus data; students can extrapolate their own translation options, explore stylistic features, investigate lexicographic aspects of translation and contribute to the development of corpus-applied translation studies.



**Corpus-branched direction**, as a subfield of corpus-applied translation studies, contributes to improving the translation process of sector-specific texts. This direction combines corpus analysis methods and approaches with the study of specific features of texts and “rules for the normative use of the translator's language” (Muñoz-Basols, Neville, Lafford, Godev 2023, 180). Corpus-branch analysis relies on a set of guidelines involving the use of text corpora, notably including texts on military topics, texts from the economic sphere, normative-legal texts, among others. The comprehensive concept of an evolutionary database (a repository of corpus-branch texts) encompasses various aspects of authentic texts representing a particular field. These texts form a unique unity of technical methods, the translator's constructive actions and convey the semantic dominants through corpus linguistics means. The inseparable connection in the parameterized database enables the analysis and study of linguistic, stylistic, terminological and other specific features of these texts. Research in specialized fields, considering the applied aspect of translation, particularly the functional, contextual-adaptation, consistent, linguistic-stylistic aspects, has practical orientation. They contribute to developing effective translation strategies and tools tailored to the specific nature of a particular field, its characteristic encoded linguistic signs and the system of recognition criteria. Research in narrow-field specialties helps identify better translation equivalents for the quality reproduction of translations within specific contexts; identify and analyze terminological vocabulary; establish branch terminological databases and glossaries that facilitate translators' understanding and use of professionally marked lexicons; consider the specifics of communicative situations, including metatranslational factors and the needs of the target audience using translations. This dichotomous division is traced back to the establishment of key translation strategies aimed not only at contextualizing words but also considering the specifics of the communicative act, effective information exchange. This allows interpreting the actual meaning of the studied speech fragment, applying a practically oriented approach, contributing to improving translators' professional training and the development of translation practice.

**Corpus-stylistic direction** aims to study the stylistic realities of translation expressed through corpus data. Researchers analyze the stylistic features of original texts and their translations, identifying stylistic similarities and differences between languages relevant for adequate translation and the author's idiolect. The proliferation of scientific texts and their significance for the development of conversational discourse requires the identification of occasionalisms in the text corpus. Interpreting conversational discourse as a corpus-linguistic phenomenon, A. Slypushova believes that conversational texts function within a specific corpus-linguistic space – the semiosphere of spontaneous speech (Slypushova 2020, 61–65). The proposition of corpus stylistics in corpus translation studies pluralizes the valuable benchmarks of text (text meaning), various aspects of speech affecting the expressiveness, effectiveness and aesthetics of translation. These units with different linguistic statuses (lexemes, terminological phrases, idioms, occasionalisms, etc.) explicate standardized speech styles in the text corpus. The richness of the text corpus with terms and scientific vocabulary accentuates specific realities. The grammatical style of speech is determined by the syntactical organization of the text, among other things, different types of sentences, word order in the sentence, the use of words belonging



to different parts of speech, sentence structure sequence representing the most common views on institutional discourse components. The perspective of reproducing different speech styles in the corpus influences the peculiarities of sentence grammatical structure. Skillfully manipulating grammatical sentence parameters, morphological, word-forming and stylistic tools of the original text, as well as equivalent variations in the recipient language, form the notion of novel corpus-stylistic translation phenomena. Various rhetorical devices (metaphors, similes, allegories, hyperboles, etc.) embodied in the linguistic form of the text corpus play an important role, providing the text with effectiveness and expressiveness, reinforcing the author's idiolect intention. As a result of the stylistic influence on the translational reception of corpus-stylistic correlation, categories of text emerged: communicativeness, informativeness, coherence, integrity, articulation, completion, referentiality, intertextuality, each represented by corresponding expressive means.

One of the tools of corpus-applied stylistics is phonostylistics, encompassing the sound formation of text along with pronunciation, intonation and rhythm (Hasselgren 2020, 143–174). The foundation of contemporary corpus phonostylistics lies in phonetic and stylistic analyses. Phonetic analysis involves studying the language's sound model and understanding its influence on perceiving and expressing the stylistic features of the text. Corpus data enable the examination of phonetic variations (positional, combinatorial, facultative phoneme variants) (Fellbaum 2019, 735–767), their transformation and equivalent representation in translation. Similarly, stylistic analysis shows speech stylistic features like lexical usage, grammatical constructions, idiomatic expressions, metaphoric phrases, etc. Therefore, through corpus data, we can identify stylistic variations in different texts and speech genres. This is why incorporating stylistics into the translation process within a text corpus seems like a logical approach, while choosing translation strategies becomes an algorithm for applying the proposed methodology.

**Corpus-lexicographic direction.** It contributes to the development of lexicographic resources – linguistic, explanatory, etymological, orthographic, orthoepic, historical, phraseological, terminological, frequency dictionaries, thesauri and glossaries – aiding translators in finding appropriate equivalents for words and expressions in the target language. Within corpus-applied translation studies, corpus lexicography encompasses the following implementation stages: collection and analysis of corpus data, comprehensively identified within textual patterns across various speech genres and styles. Lemmatization and modeling of lexical units (establishing the base form of lexical units); each lexical form of a word is assigned special tags or labels indicating its properties and contextual peculiarities. These include data such as the frequency of use of lexical units, stylistic characteristics, grammatical features, semantic nuances, and lexical translation equivalents. According to O. Zhezherun and O. Smysh, tags assist users of lexicographic resources in quickly locating needed words or phrases, understanding their meanings and contextual usage, as well as recognizing translation correspondences (Zhezherun and Smysh 2020, 71–80). Semantic analysis of lexical units, involving the establishment of their meanings, connotations, tracing the development of synonymous, antonymous,



and other semantic relationships. Isolation and examination of collocations (frequent word combinations) and phraseological expressions (fixed phrases) in corpus-applied translation studies, necessary for establishing typical word combinations in speech contexts and ensuring a more precise translation variant. Considering the stylistic features of words and expressions identified in the corpus of texts, which contextualize nuances of speech style during the creation of lexicographic resources, as well as contextual variations.

Therefore, each direction of research within corpus-applied translation studies holds significant importance in enhancing the translation process, advancing translation studies and refining translation activities. Integrating corpus methods and data into translation research and practice entails precision in presentation and unambiguous interpretation of the provided information. The discovery and expansion of functional relationships between applied linguistics, corpus linguistics, and translation studies contribute to delineating the potential of this discipline in ensuring adequate and high-quality translations across various fields of activity.

### **Conclusion**

In this paper, we demonstrated challenges that significantly impact the field of linguistic studies. These challenges transcend conventional boundaries, necessitating an adaptable approach to comprehend the complexity of language in the contemporary world. The advent of technology and globalization has fundamentally reshaped linguistic landscapes. Communication now transcends geographical constraints, extending across borders via digital platforms and online interactions. This global interconnectedness has caused the fusion of languages, the emergence of new dialects and the formation of hybrid linguistic structures. Cultural exchanges, migration, intercultural interactions have contributed to the evolution of linguistic nature, reflecting a blend of diverse linguistic elements. Furthermore, the ascendance of digital media, social networking, and the internet has revolutionized language usage and communication patterns.

Our results thus provide the theoretical foundations of studying linguistic constructs in corpus-applied translation studies which are based on understanding their theological-disciplinary nature. This involves analyzing the quantitative aspects of reproducing these constructs in a corpus of texts, contextual adequacy features, grammatical coherence, personalization of automated translation systems, pragmatic load of linguistic units. The starting point lies in reproducing linguistic patterns that demonstrate the logical and semantic completeness of lexical and grammatical features of the source language and the target language. The infological model of corpus-applied translation studies expands the variety of choices of corpus data, quantitative patterns, lexical dominants in the source and target languages, corpus-integrated translational constructs. This is results of the gnoseological determinism of linguistic units which reflect functional relationships among applied linguistics, corpus linguistics, and translation studies.



## References

1. Aharoni, Roei., Johnson, Melvin., Firat Orhan. 2019. "Massively Multilingual Neural Machine Translation." In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* 1: 3874-3884.
2. Amelia, Lolitha., Balqis, Nadira. 2023. "Changes in Communication Patterns in the Digital Age." In *ARRUS Journal of Social Sciences and Humanities* 3: 544-556. 10.35877/soshum1992.
3. Bednarek, Monika., Carr, Georgia. 2020. "Computer-assisted digital text analysis for journalism and communications research: introducing corpus linguistic techniques that do not require programming". In *Media International Australia* 181(2): 1329878X2094712. Doi:10.1177/1329878X20947124.
4. Ben-Rafael, Eliezer., Ben-Rafael, Miriam. 2015. "Linguistic landscapes in an era of multiple globalizations." In *Linguistic Landscape* 1: 19-37. Doi:10.1075/ll.1.1-2.02ben.
5. Bennett, Gena. 2010. *Using corpora in the language learning classroom: corpus linguistics for teachers*. Michigan: The University of Michigan Press: *Ann Arbor*.
6. Bisazza, Arianna, Tump, Clara. 2018. "The lazy encoder: a fine-grained analysis of the role of morphology in neural machine translation." In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018*. Association for Computational Linguistics (ACL). Brussels. P. 2871-2876.
7. Cenk, Tan. 2022. "Film Philology: The Value and Significance of Adaptation/Film Studies in Literature". In *Journal of Cultural Studies*: 31-49.
8. Crosthwaite, Peter, Ningrum, Sulistya, Schweinberger, Martin. 2022. "Research trends in corpus linguistics: A bibliometric analysis of two decades of Scopus-indexed corpus linguistics research in arts and humanities." In *International Journal of Corpus Linguistics* 28(3). DOI:10.1075/ijcl.21072.cro.
9. Crystal, David. 2023. "Philology". In *Encyclopedia Britannica*. November 29, 2023. <https://www.britannica.com/science/philology>.
10. Dash, Niladri. 2005. *Corpus linguistics and language technology: with reference to Indian languages*. New Dehli: *Mittal Publications*.
11. Demyanchuk, Yulia. 2023. "Corpus-applied translation studies as a branch of applied linguistics: a gnosiological perspective". In *Scientific journal of the Lviv State University of Life Safety, Philological Periodical of Lviv* 13(13): 51-58. <https://journal.ldubgd.edu.ua/index.php/philology/article/view/2592>
12. Errington, Joseph. 2008. "Linguistics in a Colonial World: A Story of Language, Meaning, and Power". 10.1002/9780470690765.ch4.
13. Fellbaum, Christiane. 2019. "How flexible are idioms? A corpus-based study". *Linguistics* 57(4): 735-767.
14. Geraghty, Barbara, Conacher, Jean. 2014. "Approaching migration, intercultural contact and language learning". In *Intercultural Contact, Language Learning and Migration*: 1-18.



10.13140/2.1.3172.3209.

15. Hasselgren, Angela. 2020. "Learner Corpora and Language Testing: Smallwords as Markers of Learner Fluency." In *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*: 143-174.

16. Jäger, Gerhard. 2019. "Computational historical linguistics." *Theoretical Linguistics* 45(3-4): 151-182. <https://doi.org/10.1515/tl-2019-0011>

17. Liddell, George, Scott, Robert. 2019. "φιλολογία". A Greek-English Lexicon. Perseus.tufts.edu.

18. Meyer, Charles. 2008. Pre-electronic corpora. In *Corpus linguistics: an international handbook* edited by A. Lüdeling, M. Kyto, 1: 1-14. Berlin: *Walter de Gruyter*.

19. Morgan, Hani. 2022. "Conducting a Qualitative Document Analysis." In *The Qualitative Report* 27(1): 64-77. <https://doi.org/10.46743/2160-3715/2022.5044>

20. Muñoz-Basols, Javier, Neville, Craig, Lafford, Barbara, Godev, Conception. 2023. "Potentialities of Applied Translation for Language Learning in the Era of Artificial Intelligence." In *Hispania* 106(2): 171-194. <https://doi.org/10.1353/hpn.2023.a899427>.

21. Naudé, Jacobus. 2017. "The Disciplinarity of Linguistics and Philology". In *Linguistics and Biblical Hebrew and Philology in Hebrew Studies at Society of Biblical literature*: Boston.

22. Power, Justin M., Guido, Grimm, Johann-Mattis List. 2020. "Evolutionary dynamics in the dispersal of sign languages". In *Royal Soc. Open Sci.* 7: 1-15. doi: 10.1098/rsos.191100

23. Slade, Darren. 2020. "What is the Socio-Historical Method in the Study of Religion?" In *Socio-Historical Examination of Religion and Ministry* 2(1): 1-15. 10.33929/sherm.2020.vol2.no1.01.

24. Sliepushova, Anhelina. 2020. "Father-child discourse in the animated series "Family Guy": a corpus analysis." In *Synopsis: text, context, media* 26(2): 61-65. <https://doi.org/10.28925/2311-259x.2020.2.6>

25. Stefanowitsch, Anatol. 2020. *Corpus Linguistics: A Guide to the Methodology*. Berlin: *Language Science Press*.

26. Stifter, David, Qiu Frangzhe, Aquino-Lopez, Marco. 2022. Strategies in tracing linguistic variation in a corpus of Old Irish texts (CorPH). In *Corpus studies of language through time: Special issue of the International Journal of Corpus Linguistics*: 529-553.

27. Teubert, Wolfgang, Cermakova, Anna. 2007. *Corpus linguistics: a short introduction*. London: *Bloomsbury Academic*.

28. Zhezherun, Oleksandr, Smysh, Oleh. 2020. "Automation of solving planimetry problems written in natural Ukrainian." In *Programming problems* 4: 71-80.