EXPLORING AND CONSTRUCTING PHILOSOPHICAL CONCEPTS:
A CORPUS APPROACH

Illia Ilin

Vasily Nazarovich Karazin Kharkiv National University, Philosophical Faculty,
Department of Theoretical and Practical Philosophy named after professor Johann Baptist Schad
Svoboda Sq., 4, 61022, Kharkiv, Ukraine

The article aims to summarize the tools and techniques of corpus linguistics for technologically mediated conceptual analysis in philosophy. Corpus linguistics allows for transcending subjective biases and intuitive methods, offering a departure from traditional, anthropocentric conceptual analysis with its unverified conclusions and arbitrary interpretations. Corpus-based conceptual analysis offers solutions to several challenges within contemporary social sciences, including 1) the problem of synonymy and polysemy of concepts in social sciences; 2) the problem of scientific-theoretical progress, which is hindered, in particular, by the lack of comprehensive study of theoretical developments that leads to «reinventing the wheel» in science; 3) the problem of the limitations of human capabilities and the existence of technological conditions (Big Data) and resources for effective analysis of large volumes of textual data; 4) the problem of critical reading, which imposes its interpretations on the texts, seeks hidden meanings and truths in the texts, thereby dangerously aligning itself with conspiracy thinking. Corpus-based conceptual analysis facilitates the rapid and objective identification of frequently used words, collocations, and word associations, thereby facilitating the effective and verifiable definition of concept meanings during the processing of large datasets. This approach ensures a more precise comprehension of existing concepts and the formulation of new ones. The foundations of corpus-based conceptual analysis were established in the 1960s and have since been extensively developed and applied within contemporary philosophy. Henceforth, confidence in pre-corpus conceptual research will wane, as such investigations relied on predetermined theories of philosophers, selecting fragments that corroborate those theories. This prioritization tended to favor elements in the text that aligned closely with the reader's theory, rather than those that were more frequent and thus more significant to the author's theory. Essentially, we must acknowledge a radical revolution in conceptual analysis, as technologically mediated reading enables broader empirical observations, encompassing vast volumes of texts and helps with unveiling patterns, and conceptual connections and meanings. This contrasts with the reliance on intuition and the arbitrary nature of pre-corpus research.

Key words: conceptual analysis, corpus linguistics, reading, method, philosophy, concepts.

The study and construction of concepts have long been trapped by subjective biases, intuitive approaches, unverified conclusions, and arbitrary interpretations. It is now time to break free from this captivity. The necessary transformations should not be delayed, as we already have all the means to facilitate them. We are specifically referring to conceptual analysis and the construction of new concepts based on computational methods, particularly gathered from corpus linguistics [6; 8]. This linguistic discipline allows for the identification of language patterns, generation of frequency word lists, word combinations, associations, and more [31; 32]. Consequently, it enables the provision of empirical textual data to interpreters rapidly and objectively, facilitating further theoretical work such as exploring existing concepts or creating new ones. Such data can answer researchers' questions, such as which concepts were most frequently used by a philosopher and which terms are associated with those concepts in their texts. The comprehensive

analysis of concepts directly influences the understanding of the logic, rhetoric, and politics of a philosopher's statements, as well as the understanding of his/her ability to include and give voice to all necessary actors in resolving a particular philosophical problem. Furthermore, philosophical concepts point to ways of establishing a new community with the new actors (or with new logic and politics of these actors), blurring the distinction between facts and values [19, p. 47-49; 23, p. 143-144; 22, p. 225-226; 20, p. 69, 89, 129]. Therefore, conceptual analysis is not only an exploration of the theoretical content, logic, rhetoric, and politics of philosophical discourse but also an active intervention by the thinker in the existing understanding of the nature of things, a reassembling of these things and their nature in the face of new challenges. Concept construction, accordingly, involves the formation of new heuristic theoretical statements based on existing scientific and political issues, using textual material from various philosophers. This construction allows for the identification of potential theoretical forms in previously unconceptualized logical and lexical-semantic relationships within philosophical texts. Concept construction is an essential component of conceptual analysis. The study and construction of concepts can be carried out with the aid of corpus linguistics. This article will focus on the methodology for performing such operations.

It is crucial to recognize the underlying reasons that justify the need for technologically mediated conceptual analysis. Firstly, the issue of synonymy and polysemy of concepts in the social sciences. Secondly, the problem of scientific-theoretical progress. Thirdly, the challenge posed by human limitations and the availability of technological conditions (Big Data) and resources for conducting effective analysis of textual data. Lastly, the problem of critical reading.

The first problem is related to what Fred Riggs identified as the polysemy and synonymy of concepts in the social sciences. On one hand, concepts have multiple meanings, and on the other hand, there are numerous concepts with similar meanings. However, researchers are not inclined to create neologisms for new meanings [30]. This issue significantly complicates mutual understanding and the advancement of science.

The second problem, as discussed by Michael Park, Erin Leahey, and Russell Funk [29], relates to the stagnation of scientific progress. The authors argue that the number of «disruptive» research studies has significantly declined over the past 50 years. Their study is based on the analysis of 45 million articles and 3.9 million patents. One of the proposed solutions to this problem is for scientists to read more and effectively process the vast amounts of data available in contemporary science. In other words, social scientists have largely been «reinventing the wheel» over the past half-century, partly due to inefficient analysis of existing literature.

The third problem lies in the fact that mastering vast arrays of scientific publications, even on a single topic, using only one's own, only human efforts, is a daunting and time-consuming task. The sheer volume of social communication on the Internet has led to the realization that we live in a Big Data society, which relies, among other things, on the analysis of large sets of language data. Criticizing the use of text processing technologies for reading means overlooking the fact that reading itself is a result of the technologization of human existence, or, in the terms of Bernard Stiegler, the exosomatization, the merging of the biological and the technological in human history [33, pp. 86-89]. Reading is the result of the objectification of human thinking, memory, and so on, in texts. The identification of language patterns was carried out without such technologization for at least a thousand years, but now these operations are exosomatized, i.e., taken beyond the human body into a computer program capable of assuming some of the work in the reading process. Rejecting such technologization is equivalent to marveling at the human vision in its ability to explore the cosmos but not using the «Gaia» space telescope, which tracks the movements of over 1 billion astronomical objects to create a three-dimensional map of the

I. Ilin
Вісник Львівського університету. Серія філос.-політолог. студії. 2023. Випуск 48

55

cosmos, including exoplanets. Without this exosomatic «eye», it would be impossible to determine the origin and evolution of our galaxy [7]. By relying solely on our human hearing, we could never understand what has become known thanks to the use of artificial intelligence and highly sensitive microphones, which have given birth to disciplines such as bioacoustics, capturing the acts of listening to sounds produced and perceived by plants and animals but inaccessible to the human ear (infrasound and ultrasound) [5]. Thus, without the exosomatization of our «ear», it is difficult to fully envision the struggle to protect the rights of non-human beings and the ecosystem as a whole. Numerous examples of this nature can be provided. It is clear that ignoring technologies means predisposing oneself to the repetition of the same knowledge, reducing the world to one's limited understanding, and being politically reactionary and morally irresponsible. To paraphrase a line from Christopher Nolan's film «Interstellar», reading everything may be impossible, but it is necessary.

Finally, the fourth issue pertains to critical reading. In philosophy, critical reading takes the form of symptomatic (Louis Althusser [2, p. 22]) and deconstructive reading (Jacques Derrida [11, p. 44]) among others, and it works by revealing gaps, hidden structures in texts, and/or ambivalences of language that undermine the systematicity and decenter the text. Critics operate based on dogmatized positions, intuitions, and selective fragments. As a result, they posit immutable essences and causes that cannot be verified or refuted, yet serve to define and explain all phenomena in the world and all meanings within the text [12; 24]. The explanatory model here adheres to the limited interpretation of the scholar who relies on a limited range of texts. Such reading closely aligns with politically reactionary ideas and conspiracy theories. It is precisely this mode of reading that accounts for the disregard of scientific-theoretical progress, technologically mediated reading, as well as the issue of synonymy and polysemy in scientific concepts. This mode of reading reproduces all of these problems, along with political reaction and the socio-economic mechanisms of neoliberal capitalism that hinder social progress.

Hence, it is evident that the aforementioned issues can, to some extent, be addressed through corpus-based conceptual analysis, which enables the rapid and objective processing of large textual datasets for a more precise understanding of existing concepts and the generation of new ones, all within the framework of horizontal ontology, post-critical, empirical philosophy. This philosophy follows the actors, their relations, their interpretations, and their interactions. Horizontal ontologies aspire to the ideals of ethnomethodological, empirical, verifiable, and falsifiable knowledge, as well as democratic, localized political action [21, p. 250, 252, 261]. Describing actions and connections between empirically defined objects, regardless of their objecthood (human, non-human, imaginary, real, etc.), horizontal ontologies transcend traditional ontological boundaries.

Now let us turn our attention to the overview of corpus-based conceptual research. First, let us delve into a more detailed understanding of corpus linguistics and the tools it employs. Corpus linguistics is a linguistic discipline that facilitates «the complete and systematic investigation of linguistic phenomena on the basis of linguistic corpora using concordances, collocations, and frequency lists» [32, p. 54]. Let us clarify the concepts used in this definition. A corpus is a collection of texts in digital format selected for studying the state and diversity of language. Through corpus analysis, regularities and language patterns, stable word usage, and new semantic relationships can be identified. A concordance is a compilation of all instances of word forms, each within its context. It is also used to refer to a single occurrence of a word within a context. Concordances help quickly form an understanding of the contexts in which a specific word appears and identify stable relationships with other words. Collocation refers to the consistent co-occurrence of two or more words within a short distance in a text. Collocations enable the identification of

semantically significant relations between words (idioms). A frequency list is a compilation of the most frequently occurring words in a text. Additionally, the distribution and dispersion of words in the corpus can be observed to understand the representativeness of their frequencies. A frequency list allows for an assessment of the main themes within the corpus [31, p. 9, 32, p. 170, 30; 4, p. 48-49]. Part-of-speech tagging is employed to optimize searches by identifying words based on their grammatical categories, while lemmatization brings all words to their dictionary form, i.e., the nominative case and singular number, and stemming – to reduce words to their roots. In essence, corpus linguistics is grounded in the empirically verifiable notion that word meaning is a function of word frequency and collocation, which can be quantitatively traced.

Corpus linguistics is not solely a quantitative discipline because any of its findings require interpretation. Through corpus linguistics, hypotheses, and intuitions, can be tested, refuted, or formulated for subsequent verification. Undoubtedly, corpus linguistics cannot provide absolute, objective, and exhaustive knowledge, nor does it claim to fully automate the process of interpretation [25]. However, corpus linguistics allows for empirically verifiable interpretations based on broader and more representative foundations. It does not replace the interpreter but rather expands the possibilities for interpretation, serving as an exosomatized complement to the interpreter.

To work within the framework of corpus linguistics, it is necessary to transform any electronic book into a «txt» format, clean the file from footnotes, introductions, author names, references, chapter and section headings on each page, perform lemmatization (or stemming), and then load it into a corpus analysis software such as AntConc [3]. Within this program, basic operations that are essential for conceptual analysis can be easily conducted.

Now let us turn to the techniques of corpus conceptual analysis described in contemporary literature.

Alastair McKinnon stands as the pioneering thinker who employed computer-based quantitative methodologies for analyzing philosophical texts within the context of conceptual analysis since the 1960s. His primary subject of study was the works of Danish philosopher Søren Kierkegaard. Although McKinnon produced his early works before the establishment of corpus linguistics, his methodology essentially aligns with the principles of corpus linguistics. For instance, when analyzing the term «system» in Kierkegaard's works – and hereinafter we are translating McKinnon's terminology into the language of corpus linguistics – he generated concordances containing the target term, created a mini-corpus from the text with the concordances, identified frequent collocations with the target term, and subsequently represented the relationships between other terms and the term «system» in a two-dimensional graph (where terms closer to «system» are considered more frequent collocations and vice versa) [26]. These operations enabled McKinnon to ascertain that the meaning of concepts in philosophical works depends on the co-occurrence of terms in proximity to the concepts. Exploring the frequency-based collocations of concepts allows for effective conceptual analysis and the identification of significant contexts and meanings. According to McKinnon, such a methodology enables more well-founded and verifiable interpretations of philosophical concepts.

Jean-Guy Meunier from the University of Quebec (Canada), who leads the «Laboratoire d'Analyse Cognitive de l'Information» (Laboratory for Cognitive Analysis of Information) have proposed, along with his colleagues, numerous methodological and technical solutions for conceptual analysis. We will focus on those that are closest to corpus linguistics because some algorithmic and semi-automatic techniques for conceptual analysis are more complex, require extensive programming knowledge, and have not been made freely available by these researchers. For instance, when analyzing the concept of «evolution» in the works of Henri Bergson, Meunier and his co-authors demonstrated the need to generate concordances with the canonical form of

the concept, and read through the concordances to identify other lexicalizations of the concept, such as synonyms, metaphors, hyponyms, hypernyms, and so on [10; 27; 9]. Subsequently, the identified concordances are classified based on different aspects of the concept. This classification is then subjected to categorization from the perspective of interpreter theories. Interestingly, the authors employed cluster analysis, whereby fragments of text with similar sets of terms were identified, forming the basis for assertions about the similarity of their conceptual content. However, the tools used to conduct such cluster analysis were not explicitly specified. Nonetheless, all these operations ultimately culminate in the visualization of the collocation matrix for the target term. For conceptual analysis of the concept of mind in Charles Peirce's works, Meunier utilized a random sampling technique to optimize the processing of a large number of concordances, generating only half of the entire corpus of concordances [28]. Additionally, Meunier involved experts in his research, who classified the concordances and clusters.

Mark Alfano conducted a groundbreaking investigation into Friedrich Nietzsche's moral philosophy [1]. To accomplish this, he worked with the entire corpus of the philosopher's published works. Alfano compiled a list of concepts for an examination, determining their frequency and identifying frequent collocations, taking into account potential lexicalizations (hypernym-hyponym relations). Subsequently, he collected a set of concordances, divided into smaller sections by Nietzsche. The semantic connections between frequent collocations and moral concepts across different works and the entirety of Nietzsche's oeuvre were visualized through network graphs. The mentioned concordances were then analyzed and summarized. Alfano asserts that this approach, – although not explicitly identified as such, is essentially corpus-based, – enabled the identification of empirically observable conceptual interrelationships within Nietzsche's moral philosophy. These interrelationships had gone unnoticed or were disregarded by commentators, and some assertions made by commentators in this field of study were empirically validated or falsified.

Robert Williams conducted a series of studies on the texts of W. E. B. Du Bois to investigate the relationships between concepts such as democracy and science, among others [35]. To achieve this, he employed the «regular expressions» tool (a formal language that allows for defining search rules using a set of metacharacters), the analysis of concordances and collocations, as well as *n*-grams (frequent word combinations of 2-3 and more words). The first tool enabled the identification of fragments containing the target words within a specific distance from each other. The second tool facilitated an understanding of the context in which these words were used and the connections between different texts by the scholar. The third tool highlighted terms that were associated with specific concepts. These operations enabled Williams to determine the meanings of concepts within the Du Bois corpus.

Now let us turn to our conceptual investigations to demonstrate the potential of the corpus approach. Corpus-based conceptual analysis enables the conceptualization of terms in philosophical works that were not originally conceptualized by the author but have subsequently become philosophical concepts. This has made it possible to draw attention to the functioning of relevant lexis in past philosophical works. As demonstrated in our study on Karl Marx's concept of social space, Henri Lefebvre and other researchers claimed that the German philosopher did not think about this concept [15]. Remarkably, these claims were made without even conducting a basic search for spatial lexemes in Marx's texts! However, corpus analysis revealed that in «Capital» and several other works, Marx employed words such as «space» and «spatial» in connection with formally defined concepts, thereby allowing for the construction of the concept of «social space» within Marx's framework. In this regard, the concepts that frequently occurred alongside spatial lexemes in specific fragments were identified and subsequently found in other works, thereby

58   *I. Ilin*

Вісник Львівського університету. Серія філос.-політолог. студії. 2023. Випуск 48

illustrating the evolution in the logical definition of the concept of social space in Marx's writings. Thus, instead of making unverifiable claims about the absence of the concept of social space in Marx, the corpus approach facilitated the construction of this concept based on collocates in certain fragments containing spatial lexemes. This study demonstrated that Marx possessed a distinctive understanding of social space that does not align with the concepts subsequently developed by Marxists.

Similar investigations have been conducted to conceptualize the notion of reading in the works of Francis Bacon and Hryhorii Skovoroda. Given that «reading» has become a concept in contemporary philosophy, this has provided an opportunity to empirically inquire into what these philosophers thought about reading. To achieve this, the search and generation of concordances were utilized for words related to reading, as well as their hypernyms and hyponyms (words associated with books, specific authors, and textual elements – terms, words), alongside words extracted from the identified concordances [17; 18]. The historical-philosophical context also allowed for corpus-based conceptual construction of the concept of walking in Skovoroda's texts [16].

In our view, the text does not contain hidden secrets or deep truths, as presumed in critical reading. Rather, it possesses empirically observable lexical-semantic connections that can be conceptualized based on the existence of corresponding concepts in contemporary theories. Thus, the text retains its meanings while allowing the reader to conceptualize logical and historical distance from contemporary theories by creating a new concept using empirically observable lexical-semantic connections in the corpus. It is not a matter of stretching the philosopher's conceptual content to our theoretical horizon, nor is it a demeaning act aimed at illustrating the superiority of our position. Corpus-based conceptualization enables the capturing of the meanings of specific terms and words used by various philosophers without subjugating them to the meanings of contemporary theories. Instead, it utilizes modern theories to formulate hypotheses that direct the researcher's attention toward an empirical analysis of the philosopher's corpus.

From our perspective, the assumption that a philosopher lacked a certain concept implies that contemporary scientific knowledge has formed a conceptual domain that can be used to conceptualize the relevant words in the philosopher's text (we are operating at the level of the text, without stretching or diminishing it) to engage in a dialogue between their position and the contemporary one.

Numerous similar studies can be cited, but for the sake of brevity and considering the methodological similarities among many works, – while passing the studies which utilized overly complex technologies for conceptual analysis, – let us summarize the aforementioned studies. All the mentioned works employ, among other things, low-tech corpus linguistic tools, which are readily accessible to all scholars in terms of free access to them (such as AntConc and other programs developed by Laurence Anthony), as well as the knowledge and skills required for their use. We intentionally exclude studies using shareware software, machine learning techniques, complex mathematical and statistical algorithms, or artificial intelligence.

Let us now summarize the methodological steps employed by these researchers in the context of corpus-based conceptual analysis: 1) compile a corpus of the philosopher's works (which can encompass a single publication or the entirety of their published works in digital format); 2) identify the concepts to be investigated (depending on the research objectives); 3) generate concordances (refine and expand the list of search terms, including synonyms, metaphors, hypernyms, hyponyms, and so forth); 4) read the concordances (in cases where there are a large number, generate a random sample of concordances) or concordances with frequent collocates; 5) generate clusters or *n*-grams to identify significant word combinations with the specified concepts; 6) read the concordances with clusters and collocations; 7) classify the concordances based

I. Ilin
Вісник Львівського університету. Серія філос.-політолог. студії. 2023. Випуск 48

59

on their relevance in determining the meanings of the concepts (considering the actors involved, the relationships formed within them, and how they are expected to be established in a new, expanded community); 8) summarize this classification from the perspective of the interpreter's theory; 9) visualize the collocations related to the target concept in the form of a network graph (for instance, using the freeware and user-friendly program «KH Coder» [14]). By employing this methodology, with its variations, and the use of different supplementary tools, one can achieve empirically verifiable conceptual analysis. This approach allows for the construction of concepts, a clear definition of the research boundaries, and separation of the research material from interpretations, ensuring that interpretations remain subordinate to the research material.

Thus, conceptual analysis for the first time becomes an empirical endeavor in which materials, hypotheses, and methods can be operationalized, verified, or falsified, thereby could be subject to comprehensive scrutiny. The confidence in pre-corpus conceptual studies would be deemed misplaced since such studies relied on a preconceived theory of the philosopher, selective extraction of fragments that supported that theory, and prioritized the theoretical elements in the text that aligned with the reader's theory rather than those words (concepts) that were frequent and thus significant to the author's theory. In essence, we witness a total revolution in the field of conceptual analysis as technologically mediated reading allows you to see more, explore vast textual data, and identify patterns and conceptual connections. This stands in contrast to the intuition and arbitrariness of pre-corpus investigations. The politics and morality of corpus-based conceptual analysis involve expanding one's understanding of concepts and the world, trusting the textual evidence, engaging with a broader and diverse set of actors, and testing hypotheses with empirical data. In a word, corpus-based conceptual analysis offers the possibility of discovering new conceptual horizons, and ways of connecting with the world rather than confining the world within a predetermined theory. The political relevance of employing quantitative and computational methods is poignantly discussed by Alex Williams and Nick Srnicek in «The Accelerationist Manifesto»: «We believe that any post-capitalism will require post-capitalist planning. <...> we must develop both a cognitive map of the existing system and a speculative image of the future economic system. <...> To do so, the left must take advantage of every technological and scientific advance made possible by capitalist society. <…> The tools to be found in social network analysis, agent-based modelling, big data analytics, and non-equilibrium economic models, are necessary cognitive mediators for understanding complex systems like the modern economy» [34]. It is worth noting that philosophy based on corpus-based conceptual analysis can play a crucial role in mapping contemporary society by identifying opportunities for its transformation.

One can read a text and perceive certain elements within it based on class, education, mood, and age. Each new reading of a particular book can reveal new truths and meanings to the reader. Corpus linguistics, on the other hand, acquaints us with the meanings and emphasis that the author has employed through fixed expressions and word associations. Thus, it becomes possible to discern specific themes that the author writes about. In this context, the pleasure of reading is not lost, but the text acquires its own identity that is independent of the reader. This opens up the opportunity for a genuine dialogue rather than the reader merely appropriating the text or constructing its meaning solely based on his/her own will. Corpus linguistics technologically settles the debate about whether authorial intentions can be extracted from the text by effectively demonstrating these intentions through collocations and frequency patterns in the text. Undoubtedly, with the use of new text processing methods and the expansion of the corpus size, the results of corpus analysis can be supplemented, adjusted, or rejected. However, any quantitative findings require interpretation, and the engagement with the text is shaped by hypotheses, which means that the activity of a scientist, a reader in corpus linguistics is not reduced.

Furthermore, corpus-based conceptual analysis can help overcome the critical-conspiratorial disposition of many philosophers. The issue lies in the fact that critical reading is always entwined with epistemic injustice. Miranda Fricker coined this term to describe the lack of recognition and suppression of an individual's ability to generate and share knowledge, including their personal experiences and emotions [13]. In the context of reading, this entails a persistent and irresponsible disregard for the possibility of an objective account of the frequency of words and collocations in the text in favor of one's interpretive frameworks, which selectively choose relevant fragments and impose readers' ideas onto the texts. Critics do not allow the text and its author to express themselves but instead speak strictly on their behalf, assuming that they always know more about the text than the author. Meanwhile, corpus linguistics allows us to genuinely listen to the position of the text and its author and cautiously interpret the extracted data.

References

1. Alfano M. Nietzsche's moral psychology. Cambridge University Press, 2019. 299 p.
2. Althusser L. Du Capital à la philosophie de Marx. *Lire le Capital*. Paris : Presses Universitaires de France, 1996. P. 2-79.
3. Anthony L. AntConc (Version 4.1.2) [Computer Software]. Tokyo, Japan : Waseda University, 2022. ULR: https://www.laurenceanthony.net/software (дата звернення: 27.06.2023).
4. Baker P. Using corpora in discourse analysis. Continuum, 2006. 198 p.
5. Bakker K. The Sounds of Life: How Digital Technology is Bringing Us Closer to the Worlds of Animals and Plants. Princeton University Press, 2022. 368 p.
6. Bluhm R. Corpus Analysis in Philosophy. *Evidence, Experiment and Argument in Linguistics and the Philosophy of Language*. Martin Hinton (ed.). Frankfurt am Main, Berlin, et al. : Peter Lang, 2016. P. 91-109.
7. Castelvecchi D. Mega-map of Milky Way adds depth to stars' motions. *Nature*. 2022 Jun. URL: https://www.nature.com/articles/d41586-022-01660-9 (дата звернення: 27.06.2023).
8. Caton J. N. Using linguistic corpora as a philosophical tool. *Metaphilosophy*. 2020. Vol. 51, №. 1. P. 51-70.
9. Chartrand L. et al. CoFiH: A heuristic for concept discovery in computer-assisted conceptual analysis. Proceedings of the 13th International conference on statistical analysis of textual data. 2016. Vol. 1. P. 85-95.
10. Danis J., Meunier J. G., Chartier J. F. CARCAT: Computer-Assisted Reading and Conceptual Analysis of Texts. URL: https://tinyurl.com/5a28dzwp (дата звернення: 27.06.2023).
11. Derrida J. Limited Inc a b c… . *Limited Inc*, trans. S. Weber, Evanston, IL:Northwestern University Press, 1988. P. 29-110.
12. Felski R. The limits of critique. University of Chicago Press, 2015. 240 p.
13. Fricker M. Epistemic injustice: Power and the ethics of knowing. Oxford University Press, 2007. 177 p.
14. Higuchi K. A Two-Step Approach to Quantitative Content Analysis: KH Coder Tutorial using Anne of Green Gables (Part I). *Ritsumeikan social sciences review*. 2016, Vol. 52, № 3. P. 77-91.
15. Ilin I. V. Towards a Marxian Concept of Social Space. *eidos*. 2022. №. 37. P. 44-70.
16. Ilin I. V. Approaching Hryhorii Skovoroda's ambulosophy: a corpus-based study. *Матеріали Міжнародної конференції «XXI Сковородинівські читання»*. Харків, 2022 (прийнято до друку).
17. Ilin I. V. Reading (in) Francis Bacon's «The New Organon». *«Новий Органон» (1620) та університетська філософія. До 400-річчя виходу трактату «Новий Органон» Френсіса Бекона»*. Матеріали міжнародної наукової конференції, 18-19 грудня 2020 р. Х : ХНУ імені В. Н. Каразіна, 2021. P. 58-65.

18. Ilin I. V. Reading (in) H. Skovoroda. *Матеріали Міжнародної конференції «XX Сковоро-динівські читання»*. Харків, 2021 (прийнято до друку).
19. Latour B. Facing Gaia: Eight lectures on the new climatic regime. John Wiley & Sons, 2017. 300 p.
20. Latour B. Politics of nature. Harvard University Press, 2004. 301 p.
21. Latour B. Reassembling the Social: An Introduction to Actor-Network-Theory. Oxford University Press, 2005. 301 p.
22. Latour B. The making of law: An ethnography of the Conseil d'Etat. Polity, 2010. 297 p.
23. Latour B. We have never been modern. Harvard university press, 2012. 157 p.
24. Latour B. Why has critique run out of steam? From matters of fact to matters of concern. *Critical inquiry*. 2004. No. 30, No 2. P. 225-248.
25. McEnery T., Brezina V. Fundamental Principles of Corpus Linguistics. Cambridge University Press, 2022. 275 p.
26. McKinnon A. From co-occurrences to concepts. *Computers and the Humanities*. 1977. Vol. 11. P. 147-155.
27. Meunier J. G. et al. Computer-Assisted Conceptual Analysis of Textual Data as Applied to Philosophical Corpuses. URL: https://tinyurl.com/z42rz669 (дата звернення: 27.06.2023).
28. Meunier J. G., Forest D. Computer assisted conceptual analysis (CACAT): the concept of Mind in The Collected Papers of CS Peirce. Hypothesis. 2003. URL: https://core.ac.uk/download/pdf/77617264.pdf (дата звернення: 27.06.2023).
29. Park M., Leahey E. and Funk R. J. Papers and patents are becoming less disruptive over time. *Nature*. 2023. № 613. P. 138–144. URL: https://doi.org/10.1038/s41586-022-05543-x (дата звернення: 27.06.2023).
30. Riggs F. Social science terminology: Basic problems and proposed solutions. *Terminology: applications in interdisciplinary communication*. Eds. H. B. Sonneveld, K. L. Loening. John Benjamins Publishing Company, 1993. P. 195-220.
31. Sinclair J. Corpus, concordance, collocation. Oxford University Press, USA, 1991. 179 p.
32. Stefanowitsch A. Corpus linguistics: A guide to the methodology. Language Science Press, 2020. 490 p.
33. Stiegler B. The neganthropocene. Open Humanities Press, 2018. 349 p.
34. Williams A., Srnicek N. #ACCELERATE MANIFESTO for an Accelerationist Politics. URL: https://criticallegalthinking.com/2013/05/14/accelerate-manifesto-for-an-accelerationist-politics (дата звернення: 27.06.2023).
35. Williams R. W. W. E. B. Du Bois on Democracy and Its Companion Ideas: Concordancing Their Intricacies. URL: http://www.webdubois.org/lectures/ncpsa2021.html (дата звернення: 27.06.2023).

ДОСЛІДЖЕННЯ ТА КОНСТРУЮВАННЯ ФІЛОСОФСЬКИХ ПОНЯТЬ: КОРПУСНИЙ ПІДХІД

Ілля Ільїн

Харківський національний університет імені Василя Назаровича Каразіна,
філософський факультет, кафедра теоретичної і практичної філософії
імені професора Йоганна Баптиста Шада
майдан Свободи, 4, 61022, м. Харків, Україна

Мета статті – показати інструменти та методи технологічно опосередкованого концептуального аналізу у філософії, який використовує корпусну лінгвістику. Застосування корпусної лінгвістики є необхідним для подолання суб'єктивних упереджень, інтуїтивних підходів, неверифікованих

62   *I. Ilin*

Вісник Львівського університету. Серія філос.-політолог. студії. 2023. Випуск 48

висновків та довільних інтерпретацій при традиційному, антропоцентричному концептуальному аналізі. Корпусний концептуальний аналіз зможе розв'язати декілька проблем, що існують у сучасних соціальних науках: 1) проблема синонімії та полісемії понять у соціальних науках; 2) проблема науково-теоретичного прогресу, що гальмується, зокрема, через відсутність комплексного вивчення теоретичного доробку, що приводить до «винаходження велосипеда» в науці; 3) проблема обмеженості людських можливостей та наявність технологічних умов (Big Data) та засобів для здійснення ефективного аналізу великого обсягу текстуальних даних; 4) проблема критичного читання, що накладає на текст свої інтерпретації, шукає приховані сутності, зближуючись з конспірологічним мисленням. Корпусний концептуальний аналіз дає змогу швидко, об'єктивно зафіксувати частотні слова, показати частотні поєднання та асоціації слів, що дає змогу проаналізувати значення понять під час опрацювання великих масивів даних для більш точного розуміння вже наявних понять та конструювання нових понять, при цьому працюючи в рамках горизонтальної онтології, посткритичних, емпіричних, верифікованих концептуальних досліджень. Основи корпусного концептуального аналізу було закладено ще в 1960-х рр., а в сучасній філософії вони були розроблені та застосовані в широкому колі досліджень. У статті узагальнено інструменти та методи корпусного концептуального аналізу, а також конструювання понять. Відтепер віра в докорпусні концептуальні дослідження є втраченою, оскільки такі дослідження спиралися на заздалегідь задану теорію філософа, добір потрібних для підтвердження цієї теорії фрагментів, тож у цьому процесі пріоритизувалися ті теоретичні елементи в тексті, що були найближчі для теорії читача, а не були частотними, а відтак значущими, для теорії автора. По суті, йдеться про тотальну революцію у справі концептуального аналізу, оскільки технологічно опосередковане читання дає змогу помітити більше, охопити величезні обсяги текстів, помітити патерни, концептуальні зв'язки та значення понять, що протиставляється інтуїтивізму й довільності докорпусних досліджень.

Ключові слова: концептуальний аналіз, корпусна лінгвістика, читання, метод, філософія, поняття.