

Theoretical Foundations and Specific Features of the Panchronic Corpus

TOSHTEMIROVA SITORA
Shahrisabz State Pedagogical Institute, Uzbekistan

ABSTRACT

This article analyses the theoretical foundations of the panchronic corpus, its differences from synchronic and diachronic corpora, and its distinctive features. The panchronic corpus is defined as a complex resource that, alongside the state of the language at a given point in time, enables simultaneous observation of its evolution over time, uniting historical and contemporary texts under a single search system. Also, the methodological differences between panchronic and diachronic corpora, the representativeness of the panchronic corpus, its balanced composition, such as the unique features of the unified search interface, the extended metadata system, and the role of extralinguistic tags, are explored in detail.

Keywords: Panchronia, panchronic corpus, synchronic corpus, diachronic corpus, corpus linguistics, metadata, annotation, representativeness, linguistic changes, unified search interface.

INTRODUCTION

In modern corpus linguistics, the task of systematising a collection of texts is undertaken by researchers who classify them according to a range of criteria, such as their functional objectives, structural components and chronological coverage. In researching the laws of linguistic development, determining a corpus's dynamic scope along the temporal axis (chronology) is regarded as one of the most effective methodological criteria. In

classical classifications found in the scientific literature, corpora are primarily divided into two types based on the chronological criterion: synchronic and diachronic corpora. For instance, M. Gatto distinguishes between these two types according to the corpus's chronological characteristics when categorising them (Gatto 2013). In the comprehensive classification proposed by V. V. Rikov, the chronological criterion also occupies one of the central positions. Based on this criterion, the scholar divides corpora into synchronic, diachronic, and monitor (observing the current state of the language) groups (Zakharov et al. 2021).

V. P. Zakharov, who has studied the diversity of corpus types, also lists thirteen directions for classification. In the scholar's approach, the chronological aspect is taken as a separate indicator, noting that corpora are divided into synchronic and diachronic types according to their degree of connection with the language's timeline (Zakharov 2005). Ch. Taylor, in his article 'What Is Corpus Linguistics?', also notes the division of corpora into general and special types, that special types of corpora differ according to genre, style and period, and that both types of corpus, in turn, exist in diachronic and synchronic forms.

A synchronic corpus (from the Latin *synchronus* – pertaining to the same time, coeval) represents the contemporary forms of speech units as they are used at a given moment (Abdurahmonova 2023). A synchronic corpus comprises data from a specific, limited period, aiming to depict the language's features at that time. Such corpora primarily contain texts belonging to the present day or to a particular era. The synchronic approach serves to determine the lexical composition, grammatical structure and frequency of use of language units at a given moment. The purpose of synchronic corpora is to establish empirically how the language is used today.

A diachronic corpus is a collection of texts spanning various historical periods (Daliyeva 2024). A diachronic corpus serves as the object for glottochronological and statistical analysis to determine the extent to which the language has changed over time. According to scholars such as M. Hilpert & S. Th. Gries, a diachronic corpus, as in the case of the diachronic corpus of present-day spoken English, textual resources representing

similar types of language use at successive time intervals, a feature which by its nature requires the corpus to span at least two time intervals (Ataboyev 2024). Chu-Ren Huang notes that most corpora are synchronic in nature, and that creating diachronic corpora that span a wide period is a complex process (Huang et al. 2015).

The purpose of diachronic corpora is to observe on a statistical basis the development of language units as they pass from one period to the next, their emergence or disappearance, and the acquisition of new meanings.

Both types of corpus are important parts of modern corpus linguistics, and they complement each other: synchronic corpora answer the question ‘what is happening now?’, while diachronic corpora answer ‘how did it develop and why did it change in that way?’ They answer questions such as ‘How did it develop?’ and ‘Why did it change in that way?’ As noted in the above classifications, in traditional taxonomies found in scholarly literature the chronological criterion is largely confined to two poles – synchronic and diachronic approaches. However, in some cases these corpora cannot fully reflect the continuous development of the language, its semantic and grammatical changes, and its overall picture. As a result of integrating synchronic and diachronic approaches, the concept of a panchronic corpus was developed. A panchronic corpus enabled the simultaneous observation of a language's state at a given point in time and its evolution over time. The need for a panchronic corpus is primarily determined by the different principles of lemmatisation, orthography and grammar representation in various historical and contemporary corpora (Russian National Corpus).

METHODOLOGY

Before theoretically grounding the concept of the panchronic corpus, we shall first consider the term *panchronia*, which forms the basis of this designation. The theoretical roots of the term *panchronia* can be traced back to Ferdinand de Saussure, and it was introduced into linguistics in the 1940s by representatives of

the glossematics school. The word *panchronia* comes from Greek, where “pan” means “all” and “chronos” means “time.” As a term, *panchronia* is widely used in linguistics with varying degrees of generalisation. The basic definition of this concept is as follows: ‘to consider linguistic phenomena as common (universal), timeless, and characteristic of human language in general’ (*Linguistic Encyclopedic Dictionary*). According to the concepts of V. I. Postovalova & V. Humboldt, “panchronia is generally associated with the resolution of the problem of the timeliness of the order of resolution,” such concepts as panchronia, which is usually expressed through the contradictions in dialectical motion associated with the resolution of the problem of the timeliness of time for the order of resolution, and the transition of the classical dialectical method into history, as well as synthesis, which is a present activity occurring at every moment" [Lutfullina et al. 2017].

In modern linguistics, the concept of panchronia is emerging as a paradigm that rejects the traditional dichotomy between synchrony and diachrony, studying language in its temporal integrity. As P. Lozovsky emphasised, panchronia is not merely the combination of two different methods, but the opportunity to study language independently of spatial and temporal constraints. The intersection of the synchronic and diachronic domains constitutes the methodological foundation of the panchronic approach (Łozowski 2023). The existence of two distinct interpretations of the panchronic approach requires particular attention when shaping panchronic research:

Methodological interpretation: in this, the boundary between synchrony and diachrony is considered to be variable according to the research objective.

Conceptual interpretation: in this, language is studied as an integral object along a temporal continuum, in organic connection with cognitive, cultural and experiential parameters (Łozowski 2023). These two approaches make it possible to integrate the historical and contemporary semantic layers of language units into a single system when creating the Uzbek language's panchronic corpus.

The concept of panchronia was initially proposed by F. de Saussure. The scholar viewed this approach as a set of general principles encompassing synchronia and diachronia. However, Saussure limited panchronia to abstract rules, divorced from concrete linguistic facts. According to P. Lozovsky's analysis, whereas in structuralism panchronia is seen as "abstract rules", in the cognitive-functional approach it is "functional universals" (patterns of cognitive and experiential behaviour) are manifested. P. Lozovsky distinguishes two different interpretations of panchronia. Whereas structuralism regards this concept as "abstract rules," the cognitive-functional approach explains it through "functional universals" (Łozowski 2023). In general, panchronia is the study of linguistic phenomena without reference to time. This approach also allows researchers to identify elements that remain unchanged despite the development of language over time. The theoretical perspectives outlined above demonstrate that the panchronic approach is important not only conceptually but also in practical terms. In particular, such an approach necessitates the analysis of linguistic material integrated over time within corpus linguistics. We can say that this need led to the emergence of panchronic corpora as a distinct type of corpus.

A panchronic corpus is defined as a corpus of historical texts linked to the general research of the base corpus, enabling us to trace the history of a word or grammatical structure over several centuries.

RESULTS AND DISCUSSION

Although panchronic corpora may at first glance appear similar to diachronic corpora in their external characteristics, their methodological foundations are different. The distinction between these two types of corpus is fundamental to understanding their significance and roles in linguistic research. Both types of corpus are designed to study changes in language over time. Diachronic corpora, primarily, are aimed at representing the sequential development of linguistic units over a specific period or a clearly defined historical era, helping to study

language development across historical stages. Panchronic corpora, on the other hand, allow for the analysis of linguistic phenomena across different periods, but in a unified system and within a common research framework. While in a diachronic corpus the temporal variation of language units is prioritised, in a panchronic corpus the continuity, duration and interperiodic connections of these changes are the primary focus. Diachronic corpora are an invaluable resource for linguistic research, but their scope is limited to a single clearly defined historical period. Although this type of corpus is useful for targeted research, it may create artificial barriers for researchers who wish to observe the continuous development of language evolution. A panchronic corpus, however, removes these artificial barriers, acknowledging the interconnectedness of different historical stages and providing the necessary resources and tools to study these connections. Panchronic corpora are designed to fill the gaps in diachronic corpora, providing a more complete, continuous, and comprehensive collection of data that covers specific historical periods of a language. This long-term approach is crucial for studying linguistic phenomena that unfold over centuries, such as the gradual disappearance and emergence of certain affixes or the evolution of semantic meanings. Thus, the development of panchronic corpora signifies a more holistic and integrated approach to historical linguistics. In this regard, the “Corpus Corporum” system developed by Ph. Roelli demonstrates the technical possibilities of a panchronic approach on the example of Latin (Roelli 2014).

One of the most important features of a panchronic corpus is that it enables the generalised analysis of the development of language-specific linguistic features over extended historical periods. There are no chronological boundaries between these periods. Through such a corpus, changes in a language's vocabulary, grammatical structure, phonetic and orthographic systems, as well as the development of written style and norms, can be observed consistently. As panchronic corpora bring together texts from different periods within a single system, they are an important resource for studying the historical development of languages and their adaptation to changes brought about by

social and cultural factors. This corpus can include texts from various genres and sources, such as literary works, historical documents, religious texts, and everyday writings. Therefore, a panchronic corpus serves as an effective tool for identifying similarities and differences across periods and for systematically analysing the phonetic, grammatical and semantic development of linguistic units in the diachronic process. Furthermore, the corpus's technical capabilities allow for the identification and comparison of linguistic phenomena through a single search environment.

The scientific value of a panchronic corpus and its full reflection of the research object are determined, first and foremost, by its representativeness and balanced composition. Whereas in synchronic corpora balance is achieved through the selection of texts from a specific period, in a panchronic system this concept entails a proportional distribution along the temporal axis, that is, across different historical strata. As T. McEnery and A. Hardie emphasise, a corpus, as a resource for describing language on the basis of empirical evidence, must proportionally encompass samples from all periods included in the research (McEnery et al. 2012). Otherwise, the over- or under-representation of certain periods may hinder an impartial assessment of the language's development. Therefore, in the process of forming a panchronic corpus, maintaining a balance between temporal coverage, genre diversity, and the proportion of texts is considered a distinct methodological requirement.

The availability of a single search interface is the greatest advantage of panchronic corpora. When using panchronic search, the researcher can create queries in a single interface without having to access multiple corpora several times through different interfaces, lexicon and grammatical markup formats, query languages and search restrictions (Sichinava 2024). This not only increases efficiency but also helps to use time productively. In her article, I. Y. Malisheva states that a panchronic corpus helps us to observe changes in words, including their phonetic appearance, their morphological characteristics, reflecting the word-forming variants, and emphasises that any student studying the corpus data becomes a genuine researcher, beginning to

understand the processes associated with language and semantic change (Malysheva 2023).

Another distinctive feature of the Panchronic Corpus is the extended role and complex structure of metadata, namely extralinguistic tags. In the panchronic corpus, for each text, the author, year of writing, year of publication, source type, genre, region, script (Arabic, Cyrillic, Latin), spelling variant, level of transliteration, and reliance on a critical edition, numerous tags such as manuscript copies, notes on the editorial history, and the method of digitisation are required. These tags not only facilitate searching but also ensure the scientific reliability of the analysis. Organising the linguistic annotation process on the basis of the methodological principles of international standards, particularly models such as ISO 24613-1 (Lexical Markup Framework), facilitates the integration of the panchronic corpus with other digital resources. This approach enables each lexical unit to be provided not only with morphological markers but also with special identifiers that reflect its historical variability. As a result, it becomes possible to analyse the internal dynamics of the language system on the basis of the principles of linguistic change (internal factors) proposed by U. Labov, that is, not as a random process but as a systematic evolution (Łozowski 2023). For this reason, in a panchronic corpus, extralinguistic data are regarded not merely as an “auxiliary tool” but as an integral component of the linguistic evidence. For example, when tracing the semantic evolution of a particular lexical item in historical texts, extralinguistic tags concerning the text's “genre” (poetic, prose, or official-administrative) and “source type” are of crucial importance. For a single word can have different semantic load and stylistic function in texts with different generic and regional characteristics, which in turn guarantees the objectivity of panchronic analysis.

CONCLUSION

Although the term “panhistoric corpus” is not yet widely disseminated and accepted, it accurately expresses the essence of language resources: a comprehensive, all-encompassing view of

the history of language, designed to facilitate research beyond the traditional period boundaries. The development of such corpora demonstrates the growing need to manage the complexities of historical data with more integrated and advanced methods. To date, panchronic corpora created are not merely collections of texts but sophisticated, well-developed systems that address the complexities of various orthographies, grammatical structures and annotation schemes across different periods. In theory, a panchronic corpus can link the common ancestor of various languages with one of its descendants; the oldest platform of such a corpus can also be merged with the various corpora of the language's later forms.

In conclusion, the panchronic corpus is a complex system that paints a comprehensive, all-encompassing picture of the history of language. Its creation is of immense significance not only for linguistics but also for understanding national identity and adapting cultural heritage to the digital age.

REFERENCES

1. Gatto, M. 2013. *The Web as Corpus: Theory and practice*. Cambridge University Press.
2. Zakharov, V., Mengliyev, B. & Khamroyeva, Sh. 2021. *Korpuslingvistikasi* [Corpus linguistics]. Study guide. Toshkent.
3. Zakharov, V. P. 2005. *Korpusnayalingvistika* [Corpus linguistics]. *Educational and Methodical Manual*. St. Petersburg.
4. Abdurahmonova, N. 2023. *Korpuslingvistikasi* [Corpus linguistics]. Textbook. Toshkent.
5. Daliyeva, M. 2024. *Diachronic Corpora and Language Evolution Over Time*. Web of Teachers.
6. Ataboyev, N. 2024. *Mediamatnlardiaxronikkorpusidatilirivojini gempiriktahlitamoyillari* [Empirical analysis principles of language development in the diachronic corpus of media texts]. Doctoral dissertation abstract, Fergana.
7. Huang, C.-R. & Yao, Y. 2015. Corpus linguistics. In *International Encyclopedia of the Social & Behavioral Sciences*. 2nd Ed. Elsevier.
8. Russian National Corpus. n.d. Panchronic corpus: integration of historical and modern corpus resources. Available online: <<https://ruscorpora.ru/corpus/panchron>>.

9. Linguistic Encyclopedic Dictionary. n.d. Retrieved from <<https://rus-lingvistics-dict.slovaronline.com>>.
10. Lutfullina, G. F. & Khayrutdinova, G. A. (2017). Yavleniepanhronii v funkcionirovaniivremennyh form russkogoyazyka [The phenomenon of panchrony in the functioning of tense forms of the Russian language]. *Vestnik Vyatskogo Gosudarstvennogo Universiteta*, 5, 52-54.
11. Łozowski, P. 2023. In search of panchrony: Saussure versus cognitive linguistics. *Lublin Studies in Modern Languages and Literature*, 47/2, 153-164. <https://doi.org/10.17951/lsmll.2023.47.2.153-164>.
12. Roelli, Ph. 2014. The corpus corporum, a new open Latin text repository and tool. *Archivum Latinitatis MediiAevi*, 72, 289-304.
13. McEnery, T. & Hardie, A. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge University Press.
14. Sichinava, D. V. 2024. Panchronicheskijkorpus: integraciya istoricheskiiisovremennyhkorpussyhresursov [Panchronic corpus: integration of historical and modern corpus resources]. *Trudy Instituta Russkogo Yazykaim*, 2, 336-353.
15. Malysheva, I. Yu. 2023. Ispolzovaniedannyh Nacionalnogokorpus arusskogoyazykapriizucheniileksiki v shkole [Using data from the Russian National Corpus when studying vocabulary at school]. *Obrazovaniei Pravo*, 4, 100-111.

TOSHTEMIROVA SITORA

PHD STUDENT, UZBEK LANGUAGE,
SHAHRISABZ STATE PEDAGOGICAL INSTITUTE,
UZBEKISTAN.
E-MAIL: <SHERARABOV4@GMAIL.COM>