Sciental Journal of Education Humanities and Social Sciences, Vol. 4. 5, (2026)
ISSN Online: 2961-0389
***Website**: https://scientaljournals.com/index.php/SJEHSS/index*

Submitted: March 22, 2026/ Accepted May 11, 2026/ Published May 21, 2026

DESIGNING A POETIC CORPUS INTERFACE: NLP APPROACHES

Nigora Tursunova Komil qizi
Lecturer, National University of Uzbekistan
E-mail: nigora.tursunova1128@gmail.com
Phone: +998879424999

Abstract:
This paper provides a comprehensive analysis of user interface design in poetic corpora based on approaches from computational linguistics and natural language processing (NLP). The relevance of the study lies in the need to develop intuitive and functional interfaces for efficient analysis and use of poetic texts in digital environments. The study examines the key components of interfaces designed for poetic corpora, their role in linguistic-poetic analysis, and their impact on user experience. Methodologically, the research employs semantic annotation techniques, automated analysis algorithms, and corpus linguistics methods. The findings indicate that a well-designed interface significantly optimizes the processes of search, analysis, and interpretation for users. Additionally, the paper proposes practical recommendations for improving user interfaces in poetic corpora.

Keywords: Poetic corpus, user interface, NLP, semantic annotation, linguopoetics, usability.

INTRODUCTION

Within the framework of modern computational linguistics and corpus linguistics, the user interface (User Interface) occupies a central place in the processes of processing, analyzing, and modeling texts in digital environments. The interface manifests not merely as a technical tool, but as a cognitive and functional platform that organizes the interaction between humans and linguistic data. Poetic texts, in particular, differ fundamentally from ordinary prose texts in their multi-layered semantic structure, rhythmic-metrical organization, and aesthetic charge. For this reason, the need arises to develop specialized, flexible, and intelligent interface solutions for organizing such texts in corpus form and for their effective use.

Poetic corpora constitute a complex scholarly system that forms at the intersection of linguopoetics, stylistics, corpus linguistics, and the digital humanities. In this system, the interface does not merely serve as a technical component, but is regarded as a methodological tool that defines research strategy, guides user activity, and serves the interpretation of linguistic analysis results. Through the interface, the user formulates search parameters, selects annotated units, observes contextual connections, and gains the capacity for complex analysis of poetic structures.

From this perspective, modern poetic corpus interfaces are required not to be limited to simple search and result-display functions, but to be integrated with advanced NLP (Natural Language Processing) technologies such as semantic annotation, metrical analysis, rhyme scheme

Sciental Journal of Education Humanities and Social Sciences, Vol. 4. 5, (2026)
ISSN Online: 2961-0389
***Website**: https://scientaljournals.com/index.php/SJEHSS/index*

Submitted: March 22, 2026/ Accepted May 11, 2026/ Published May 21, 2026

identification, automatic extraction of stylistic units, and contextual visualization. This makes it possible to interpret the interface not merely as a tool for use, but as an integral part of the scientific cognitive process.

LITERATURE REVIEW

Within the framework of Human–Computer Interaction (HCI) theory, the interface is interpreted as an important mediating layer that organizes the mutual dialogue between humans and machines. According to the HCI approach, the interface is not merely a channel for transmitting information, but a complex system that shapes the user's cognitive activity, decision-making process, and strategy of interaction with the system. In this field, the principles of "user-centered design" advanced by Ben Shneiderman, as well as the concept of "usability" developed by Don Norman, are widely applied as a theoretical and practical foundation for interface design. These approaches serve to improve the quality of human–computer interaction by making the interface intuitive, consistent, and effective.

Within corpus linguistics, the interface carries particular methodological significance, since it directly determines the researcher's strategy for working with linguistic data. Through the interface, the user identifies search parameters, sorts linguistic units, observes contextual connections, and interprets statistical results. For this reason, it is appropriate to evaluate the interface not merely as a technical tool, but as a cognitive-instrumental system that governs and directs scholarly analysis.

Modern natural language processing (NLP) systems are significantly expanding the possibilities for automating complex linguistic operations through the interface. In particular, processes such as morphological analysis, syntactic parsing, semantic annotation, and discourse analysis are carried out in an integrated manner with the interface. As a result, the user gains the opportunity to conduct rapid and in-depth scholarly analysis of large-scale texts. This demonstrates the necessity of interpreting the interface as an integral, highly functional component of computational linguistics and corpus research.

METHODOLOGY

In this study, a complex scholarly approach was employed for the purpose of investigating the poetic corpus interface and identifying its linguistic and functional capabilities. In the research process, an integrated system of methods from computational linguistics, corpus linguistics, and natural language processing (NLP) was applied.

In particular, the following scholarly-methodological approaches were taken as a basis:

— **Corpus-based analysis** — corpus linguistics methods were applied to identify the structural, lexical, and statistical features of poetic texts. This approach made it possible to identify recurring units, collocations, and poetic constructions in a large-scale text database.

— **Semantic annotation** — was aimed at semantically tagging linguistic units within the text and identifying their contextual functions. In the annotation process, the aesthetic charge, degree of figurativeness, and semantic relations of poetic units were analyzed.

Sciental Journal of Education Humanities and Social Sciences, Vol. 4. 5, (2026)
ISSN Online: 2961-0389
***Website**: https://scientaljournals.com/index.php/SJEHSS/index*

Submitted: March 22, 2026/ Accepted May 11, 2026/ Published May 21, 2026

— **Comparative study of interface design principles** — existing corpus interfaces were compared on the basis of criteria of ergonomics, usability, intuitiveness, and functionality. This served to draw theoretical and practical conclusions for improving the poetic corpus interface.

— **Conceptual analysis of NLP algorithms** — the degree of adaptability of natural language processing algorithms to poetic texts, and their possibilities in metrical, phonetic, and semantic analysis, were evaluated from a theoretical standpoint. In particular, the integration of automatic annotation, search, and analysis modules with the interface was studied.

As research material, poetic texts in Uzbek and other languages were selected, along with corpus systems in practical use and their interfaces. On the basis of the selected materials, the influence of the interface on the linguistic analysis process, its functional effectiveness, and its impact on user experience were studied in a comprehensive manner.

RESULTS

Features of the poetic corpus interface

As a result of the conducted research, it was established that the poetic corpus interface differs significantly from linguistic corpus interfaces in functional and structural terms. This difference is explained, first and foremost, by the distinctive linguopoetic nature of the poetic text — its rhythmic-metrical organization, phonetic harmony, and aesthetic charge. For this reason, the poetic corpus interface must possess multi-layered and specially adapted functional capabilities that differ from ordinary text search systems.

Analyses demonstrated that one of the main features of the poetic corpus interface is the organization of the search process on the basis of poetic units. In particular, the possibility of conducting searches based on line (*misra*), stanza, meter (metrical model), rhyme scheme, and rhythmic structure significantly increases the scholarly value of the interface. This approach enables the user to identify and analyze not only lexical units, but also poetic constructions.

The poetic corpus interface is also distinguished by its provision of contextually extended search. That is, the user gains the possibility of viewing a selected unit not only within an individual line, but also within the context of the entire stanza, poem, or authorial corpus. This serves a deeper understanding of the functional and aesthetic role of poetic units.

According to the results, modern poetic corpus interfaces are characterized by the following principal functional features:

— support for metrical and rhythmic analysis;
— the capability of identifying rhyme and phonetic harmony;
— a multi-level (line, stanza, text) search system;
— the extraction of semantic and stylistic units;
— the visual and interactive presentation of results.

Furthermore, it was established that interfaces integrated with NLP technologies create the possibility of automatically identifying and classifying poetic units, significantly accelerating the analysis process. This demonstrates that the poetic corpus interface is of importance not only from the standpoint of convenience, but also scholarly effectiveness.

Sciental Journal of Education Humanities and Social Sciences, Vol. 4. 5, (2026)
ISSN Online: 2961-0389
***Website**: https://scientaljournals.com/index.php/SJEHSS/index*

Submitted: March 22, 2026/ Accepted May 11, 2026/ Published May 21, 2026

Interface components

The poetic corpus interface is a functionally and structurally complex multi-layered system that operates on the basis of the mutual integration of three principal components: **visual, functional, and semantic components**. These components, in their interrelated state, ensure effective interaction between the user and poetic data.

— **The visual component** serves to represent the poetic text graphically and structurally. Through this component, the line-by-line structure of the poetic text, stanza divisions, and rhythmic and metrical units are presented visually. As a result, the user perceives the aesthetic and structural features of the text in a rapid and intuitive manner.

— **The functional component** enables the user to perform operations within the corpus such as search, filtering, sorting, and sub-corpus creation. This component defines the practical effectiveness of the interface and optimizes working with large-scale poetic data.

— **The semantic component** is directed at identifying the semantic relations, contextual connections, and stylistic features of linguistic units. This component plays the principal role in carrying out a deep linguopoetic analysis of the poetic text.

The integration of these three components transforms the poetic corpus interface from merely a technical tool into a complex scholarly analysis platform.

NLP-based capabilities

The development of modern natural language processing (NLP) technologies has significantly expanded the functional capabilities of the poetic corpus interface. The integration of NLP algorithms with the interface creates the possibility of automatically analyzing and structurally processing poetic texts.

In particular, the following directions are of special importance:

— **Automatic rhyme identification** — enables the algorithmic identification of phonetic harmony and rhyme schemes in poetic texts;

— **Metrical analysis algorithms** — carries out a formal analysis of poetic structure by automatically identifying meter, rhythm, and syllabic organization;

— **Identification of semantic connections** — serves to identify the semantic relations between lexical units in the text and to express them in the form of semantic networks;

— **Contextual search systems** — enables the execution of searches based not only on a keyword, but on context and semantic proximity.

As a result, NLP technologies transform the poetic corpus interface from a conventional search system into an intellectual analysis platform, ensuring that linguopoetic research is carried out in an automated and deepened form.

DISCUSSION

The poetic corpus interface, within the framework of modern computational linguistics, corpus linguistics, and the digital humanities, is regarded not merely as a technical instrument, but as an important methodological component of the scholarly cognitive process. It serves as a mediating layer between the researcher and linguistic data, defining the conceptual and

ISSN Online: 2961-0389
***Website**: https://scientaljournals.com/index.php/SJEHSS/index*

Submitted: March 22, 2026/ Accepted May 11, 2026/ Published May 21, 2026

operational model for accessing, sorting, and analyzing information. From this perspective, the interface manifests not only as an organizational tool of scholarly analysis, but also as a cognitive-analytical environment that shapes research strategy.

Through the interface, the user (researcher) constructs a search strategy not merely as a technical query, but as an analytical process aimed at testing linguistic hypotheses. The process of defining search parameters, selecting corpus segments, and working with filters and metadata becomes part of the research design. At the same time, the interpretation of results obtained is based not only on statistical outputs, but also on the contextual, semantic, and visual indicators provided by the interface. This transforms the interface into an active tool in the formation of scholarly conclusions.

User-Centered Design (UCD) principles constitute one of the main methodological approaches that determine the effectiveness and scholarly value of the poetic corpus interface. The UCD approach presupposes adapting the interface to the user's cognitive load, linguistic competence, and research objectives. Within this framework, ergonomic and functional indicators such as usability, intuitiveness, ease of learning (learnability), error tolerance, and flexibility are regarded as important scholarly indicators.

In particular, the usability indicator is a factor that directly influences the scholarly research process, serving to minimize the cognitive and operational resources expended by the user in interaction with the system. Intuitiveness, for its part, ensures that the interface structure is close to natural logic, thereby accelerating the researcher's process of mastering the system. Flexibility enables the configuration of the interface for various scholarly tasks — operations such as corpus analysis, semantic searching, and identifying poetic structures. For this reason, the design of the poetic corpus interface is evaluated not merely as a matter of software engineering, but as a complex scholarly problem requiring the integration of cognitive linguistics, corpus methodology, and NLP technologies. The excellence of the interface is determined not only by its technical capabilities, but also by the degree to which it supports scholarly analysis.

CONCLUSION

Designing the poetic corpus interface constitutes an important scholarly-practical field that forms at the intersection of modern computational linguistics, corpus linguistics, and natural language processing (NLP). The interface serves not merely as a technical tool, but as an integral component of the processes of processing linguistic data, conducting semantic analysis, and modeling poetic structures. For this reason, it is one of the principal factors determining the quality of data access, the accuracy of analysis, and the level of interpretation in corpus-based research.

The analyses conducted demonstrate that the development of modern NLP technologies (including automatic morphological analysis, syntactic parsing, named entity recognition, and semantic embedding models) is significantly expanding the functional capabilities of poetic corpus interfaces. With the aid of these technologies, the possibility of automatically identifying and visualizing metrical structure, rhyme schemes, lexical-semantic fields, and

ISSN Online: 2961-0389
***Website**: https://scientaljournals.com/index.php/SJEHSS/index*

Submitted: March 22, 2026/ Accepted May 11, 2026/ Published May 21, 2026

stylistic units in poetic texts is emerging. As a result, the interface is evolving not merely into a tool that supports scholarly analysis, but into an intellectual platform that automates it.

At the same time, large-scale systems developed in the field of corpus linguistics — for example, projects such as Sketch Engine and the International Corpus of English (ICE) — are confirming in a practical sense the role of the interface in scholarly analysis. The experience of these systems demonstrates that the intuitiveness, flexibility, and semantic richness of the interface directly influences the accuracy and speed of research results.

In the future, the development of specialized interfaces for poetic corpora in the Uzbek language will remain one of the pressing scholarly directions. In this context, the creation of semantically annotated corpora, the development of automatic poetic analysis modules based on NLP, and the introduction of user-centered adaptive interfaces are defined as important tasks. In particular, the creation of intelligent corpus systems integrating metrical analysis and rhyme identification algorithms for Uzbek poetry may ensure a new stage in linguistic research.

In general, the development of the poetic corpus interface is of not only technological, but also theoretical-methodological significance, and will occupy an important place in the future progress of Uzbek linguistics and the digital humanities.

REFERENCES

1. Norman, D. A. (2013). The Design of Everyday Things. Revised and Expanded Edition. Basic Books.
2. Shneiderman, B., Plaisant, C. (2016). Designing the User Interface: Strategies for Effective Human–Computer Interaction. Pearson.
3. Jurafsky, D., Martin, J. H. (2020). Speech and Language Processing (3rd ed. draft). Stanford University.
4. Manning, C. D., Schütze, H. (1999). Foundations of Statistical Natural Language Processing. MIT Press.
5. McEnery, T., Hardie, A. (2012). Corpus Linguistics: Method, Theory and Practice. Cambridge University Press.
6. Preece, J., Rogers, Y., Sharp, H. (2015). Interaction Design: Beyond Human–Computer Interaction. Wiley.
7. Dix, A., Finlay, J., Abowd, G. D., Beale, R. (2004). Human–Computer Interaction. Pearson/Prentice Hall.
8. Nielsen, J. (1994). Usability Engineering. Morgan Kaufmann.
9. Baecker, R., Grudin, J., Buxton, W., Greenberg, S. (1995). Readings in Human–Computer Interaction. Morgan Kaufmann.
10. Bird, S., Klein, E., Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media.
11. TEI Consortium. (2023). TEI Guidelines: Text Encoding Initiative Guidelines. https://tei-c.org
12. Sinclair, J. (1991). Corpus, Concordance, Collocation. Oxford University Press.
13. Stubbs, M. (2001). Words and Phrases: Corpus Studies of Lexical Semantics. Blackwell.

ISSN Online: 2961-0389
***Website**: https://scientaljournals.com/index.php/SJEHSS/index*

Submitted: March 22, 2026/ Accepted May 11, 2026/ Published May 21, 2026

1. McCarthy, M., O'Keeffe, A. (2010). The Routledge Handbook of Corpus Linguistics. Routledge.
2. Baker, P. (2006). Using Corpora in Discourse Analysis. Continuum.
3. Russian National Corpus (RNC) scholarly and methodological works. https://ruscorpora.ru
4. Meyer, C. F. (2002). English Corpus Linguistics: An Introduction. Cambridge University Press.