DOI: https://doi.org/10.31392/NPU-nc.series9.2024.27.08
UDC: 81'373.421:003.071

Hamest S. Tamrazyan
PhD in Philology, Associate Professor,
Department of Germanic Languages,
Faculty of Foreign Philology,
Mykhailo Dragomanov State University of Ukraine,
Kyiv, Ukraine;
Researcher,
Laboratory of Digital Humanities,
College of Digital Humanities,
Federal Polytechnic Institute of Lausanne (Switzerland)
https://orcid.org/0000-0001-9630-0195
e-mail: a.s.tamrazyan@udu.edu.ua; hamest.tamrazyan@epfl.ch

DESCRIPTIVE TERMINOLOGY IN UKRAINIAN EPIGRAPHY: A CORPUS-BASED STUDY OF SYNONYMOUS TRIPLET "ГРАФІТІ", "НАПИС", AND "ЗАПИС"

Bibliographic Description:

Tamrazyan, H. (2024). Descriptive Terminology in Ukrainian Epigraphy: a Corpus-Based Study of Synonymous Triplet "графіті", "напис", and "запис". *Journal of Mykhailo Dragomanov State University of Ukraine. Series 9. Current Trends in Language Development*, 27, 92–110. https://doi.org/10.31392/NPU-nc.series9.2024.27.08

Abstract

The main goal of our study is to analyse the descriptive terminology used in academic discourse dedicated to Ukrainian epigraphy employing a corpus-based approach. Using the BootCat tool and custom tuples such as "епіграфіка графіті напис" and "корпус Україна графіті" we scraped the Web and generated a sample corpus containing Ukrainian texts related to Epigraphy. Our goal was to analyse the variety of epigraphic terms circulating the digital domain and the peculiarities of their usage. Our findings reveal variations in term usage in Ukrainian epigraphic texts, offering insights into their academic application. The case study showcased how synonyms can not only enrich a language, but also impose challenges in standardising terminology due to linguistic variations and the influence of individual scholars. The findings emphasise the need for a nuanced approach to developing standardised vocabularies, considering the linguistic, historical and cultural contexts. This study serves as a valuable example of how digital tools can aid in understanding and addressing the complexities inherent in preserving and studying cultural heritage in the digital age.

Keywords: *harmonised / standardised terminology, SKOS, digital epigraphy, corpus linguistics.*

1. Introduction.

Epigraphy is a discipline dedicated to the study of inscriptions which represent both tangible and intangible aspects of cultural heritage. They are especially valuable for deciphering historical and cultural narratives of any nation and country (*ЕПІГРАФІКА*, n.d.). These artefacts help scholars gain insights into the historical events and figures,

community life, personalities, mentality, lifestyle, language, habits, beliefs, skills typical of the time and practice of community. As such, these speaking stones not only enrich our understanding of the historical timeline, but also help us appreciate the complexities of human civilizations across different historical periods.

Central to the discipline of epigraphy, as to any other academic discipline, is the significance of terminology. The shared understanding and precise use of terms across the scientific discourse is crucial for the accurate interpretation and classification of inscriptions, which is essential for a coherent understanding of historical and cultural narratives. However, the challenges posed by linguistic variations are formidable, necessitating a deep understanding of these terms within their cultural contexts.

2. Literature Review.

The digital revolution has brought about a transformative impact on the preservation and research of cultural heritage (*A New Companion to Digital Humanities*, 2015). Advancements in digital technologies have democratised access to epigraphic collections (Elliott, 2014) and introduced innovative methodologies for their preservation, analysis and decipherment (Lerones & Llamas, 2021).

The European Union has been committed to digitising cultural heritage for a long period of time (Miric, 2016), (Dobreva, 2005), (Marinković & Maslikovic, 2019). This is evident in the various policies[1], project[2] initiated in Europe. According to "The 2021 European Commission Report[3]" digitisation, online accessibility, and digital preservation of cultural heritage have made significant progress during the last decade. However, challenges such as the need for standardised metadata, improved data quality, and attention to language and terminology remain. In this context, the importance of standardised terminology and the need for a systematic approach to terminology in epigraphic studies is evident. Here, special attention should be paid to the role of synonyms and related terms that enriches the language, yet they also introduce a potential for ambiguity and confusion (*European Commission Report on Cultural Heritage*, 2019).

Our research strives to compare the epigraphic terminology used in English, Ukrainian, and Armenian contexts, aiming to shed light on linguistic and cultural differences in epigraphic terminology. By leveraging digital tools and corpus-based analysis, this comparative analysis seeks to contribute to the development of standardised terminology, thus filling a gap in digital epigraphy in Ukraine and Armenia. Specifically, this article presents a case study of one set of synonymous terms in Ukrainian, offering a detailed analysis of the nuances and applications of these terms within the epigraphic domain.

The significance of this study lies in its potential contributions to the broader discourse on digital humanities, digital epigraphy (Soriano & Espinosa, 2021) and Simple

---

[1] 2015-2017 [https://digital-strategy.ec.europa.eu/en/library/european-commission-report-cultural-heritage-digitisation-online-accessibility-and-digital]
EUROPEAN COMMISSION (2021): Recovery plan for Europe https://ec.europa.eu/info/strategy/recovery-plan-europe_en].
Strategy 2020-2025 Empowering Digital Change
https://pro.europeana.eu/page/strategy-2020-2025-summary#:~:text=Empowering%20digital%20change.&text=This%20vision%20for%20Europeana%20imagines,a%20sense%20of%20European%20identity.

[2] https://pro.europeana.eu
https://research-and-innovation.ec.europa.eu/research-area/social-sciences-and-humanities/cultural-heritage-and-cultural-and-creative-industries-ccis/cultural-heritage-cloud_en
https://eureka3d.eu

[3] EUROPEAN COMMISSION (2021): Europe's Digital Decade: digitaltargets for 2030 [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-Gt-digital-age/europes-digital-decade-digital-targets-2030_en]

Knowledge Organisation Systems (SKOS)[1]. This research aspires to standardise epigraphic terminology (Mugridge, 2007) in Ukrainian and map it with existing English harmonised/standardised vocabularies(Bates & Maack, 2009), thereby facilitating the preservation of epigraphic cultural heritage and advancing the academic discipline of digital humanities(Heřmánková et al., 2022).

3. Aims and Objectives.

The aim of this study is to analyse the linguistic and cultural nuances of a selected set of synonyms in Ukrainian epigraphic terminology. This will contribute to the final goal of our research which is the creation of SKOS for the epigraphic domain, the first step of the process being the harmonisation/standardisation of terminology (Mugridge, 2007) within the field. The immediate goal of this article is to analyse the set of synonymous terms, which nominate the main subjects of research of Ukrainian epigraphy. This analysis requires a detailed examination of the usage patterns of these terms within our corpus, focusing on their frequency, context, and the variations in which they appear. Our goal is to understand the roles and nuances of these terms in academic discourse and their significance within the field. Moreover, we aim to understand how the linguistic and cultural contexts influence their usage in the epigraphic studies and uncover the reasons behind the preferences for specific terms over others. The analysis of all these aspects will help to shed light into the linguistic, cultural or other considerations that determine the terminology choices in Ukrainian and Armenian epigraphy.

Based on the insights gained from the analysis, we aim to suggest clear usage patterns for the selected synonyms to reduce ambiguity and to enhance the clarity and coherence of epigraphic discourse.

Lastly, we aim to highlight the role of digital tools and corpus-based analysis, and the efficiency and potential of these methodologies in digital humanities in general and digital epigraphy in this particular case. By emphasising the application of these innovative methodologies, the study encourages their broader adoption and application in the field. Through these objectives, our research strives to make a meaningful contribution to the field of digital epigraphy by introducing a more standardised terminology that is informed by both linguistic insights and cultural understanding.

4. Methodology.

The methodology of this study is designed to analyse the use of a specific set of synonymous terms in Ukrainian epigraphy – “графіті”, “напис”, and “запис”, utilising a combination of digital tools and corpus-based analysis to achieve comprehensive insights into their application and implications within the field.

The objective of this methodological approach is to provide a thorough comparative analysis of these epigraphic terms based on quantitative and qualitative data. As Karniienko noted (Korniienko, 2010a), these terms have been employed as synonyms to enhance the book’s language and avoid tautology. However, when compiling a SKOS for epigraphy, it is crucial to conduct a detailed analysis of the synonymous set to identify the preferred label for each concept precisely. Additionally, understanding how to represent the semantic relationship between synonyms is vital for accurately determining their connection, whether as alternative labels or related concepts.

Corpus Compilation and Selection Criteria:

The language data for our research is the specialised sample corpus containing Ukrainian files specifically related to Epigraphy and inscriptions. This corpus was collected

[1] https://www.w3.org/TR/skos-reference/

using the BootCat tool, employing custom tuples ***"епіграфіка графіті напис"*** and ***"корпус Україна графіті"*** by scraping the web for relevant texts systematically. Our goal was to survey the range of epigraphic terms circulating on the internet.

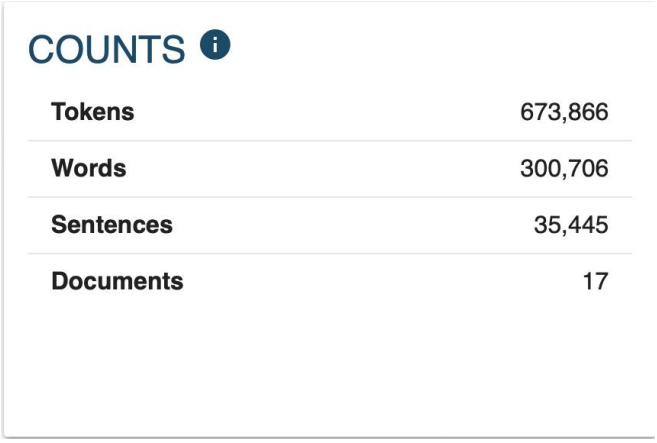


Figure 1: The counts of the corpus

The rationale behind the choice of the terms is as follows: first, the term *"епіграфіка"* is the main term that represents this academic discipline we are interested in. Second, terms *"графіті"* and *"напис"* are the most frequently used terms according to preliminary calculations in published editions dedicated to Ukrainian epigraphic heritage. Moreover, these terms represent the object of research of the epigraphic discipline. Third, the term *"корпус"* is selected as it gives the opportunity to include in our corpus not only scientific articles, but also special published editions which particularly represent the corpus of Ukrainian epigraphy. And finally, the term *"Україна"* is used to restrict the texts dedicated to the epigraphic heritage of Ukraine.

As a result, the compiled corpus comprises academic publications, including articles, dissertations, and web resources dedicated primarily to Ukrainian epigraphic heritage, providing a rich dataset for analysis.

As expected, the majority of the resources included in our corpus turned out to be primarily academic publications, providing a valuable insight into the academic discourse surrounding these terms within the field of Epigraphy.

A key limitation of our sample corpus is the substantial inclusion of texts by Korniienko V. V. on the epigraphic heritage of Saint Sophia's Cathedral in Kyiv (Korniienko, 2010a). This led to an over-representation of what could be termed "the personal academic vocabulary" of a single author. However, it is noteworthy that Korniienko V. V. acknowledges the collaborative nature of the corpus on Sofiiski graffiti, particularly under the guidance of Dr. N. Nikitenko, who contributes significantly to the study of epigraphic monuments from the old Russian period *"Над складанням корпусу софійських графіті працює не лише автор книги. Вивчення епіграфічних пам'яток здійснюється під керівництвом доктора історичних наук, професора Н. Нікітенко, яка бере особисту участь у вивченні ряду написів давньоруського часу"* (Korniienko, 2010a).

Analysis of Linguistic Variations and Synonym Usage:

The core of our methodology revolves around the detailed analysis of the selected set of synonyms, focusing on their frequency, context, and variations in usage within the

Ukrainian epigraphic discourse. Using LancsBox[1] and SketchEngine2, software tools developed for corpus analysis, we conduct collocation analysis to identify patterns of usage and associations between these synonyms and other terms within the corpus. This approach enables us to bring to light the nuances in different usage patterns of the synonymous set, including the contexts in which each term is preferred and the potential reasons behind these preferences.

The structure of the publications dedicated to epigraphy, especially the structure of the nomenclature of each inscription in the initial chapters of Korniienko's works, greatly influenced our analysis. This structure, listing each inscription's name, ID, dating, and type[3]revealed that one of our terms of interest, *"графіті"* was used as a category label containing information about the ID of each inscription, resulting in its frequent repetition in the corpus. The repetitive usage primarily as a category label led to a skewed representation of the term's frequency and did not reflect its diverse or contextual use within broader academic discourse.

I.2. *Фреска з образом невідомої святої мучениці*

Знаходиться в арці-переході між західними компартиментами центральної нави та приділа св. Іоакима та Анни на східній грані лопатки стіни (табл. II, 1). Поверхня стіни пошкоджена численними дрібними вибоями та подряпинами, що нині закриті шпаклівкою, місцями стародавній тиньк слабо розчищений від пізнішої олійної фарби XIX ст. В нижній частині фреска має втрати. На ній виявлено сім графіті, яким надано порядкові номери від 2095-го по 2101-й (табл. II, 2).

Графіті № 2095 (табл. IV, 1)

Складові напису пошкоджені вибоями, слабо розчищені від шпаклівки та пізнішої олійної фарби XIX ст., тиньк навколо зберігся погано. Втім, можемо встановити, що текст з самого початку містив лише п'ять знаків, які можуть бути прочитані як ÅAP05.

Датування. За формою написання c з язичком догори графіті найвірогідніше датувати початком XIV – кінцем XV ст.

Категорія: абетки, склади та окремі літери.

Тиньк в місці розташування напису пошкоджений та слабо розчищений від олійної фарби XIX ст. На початку рядка відносно впевнено читається літера Å, на деякій відстані від неї – 0 та A.

Датування. За формою прорізів час появи графіті може бути окреслений другою чвертю XI – кінцем XVI ст.

Категорія: фрагмент напису.

Графіті № 2097 (табл. IV, 3)

На поверхні стіни прокреслений напис, що складається з чотирьох літер.
Дві перші впевнено визначаються як АБ, дві останні, менш впевнено, як 0Б.

Датування. За формою написання ь з округлою петлею графіті датується в межах другої чверті XI – кінця XII ст.

Категорія: абетки, склади та окремі літери. (Korniienko, 2010a)

To address this challenge, we refined our search criteria to exclude instances where *"графіті"* was immediately followed by the sign ***"№"*** or a numeral. This adjustment

[1] http://corpora.lancs.ac.uk/lancsbox/
[2] https://app.sketchengine.eu/
[3] It is worth mentioning that the term ***"Категорія"*** as used in this context, likely aligns with the "Type of Inscription" found in the EAGLE vocabulary and other ontologies dedicated to Epigraphy.

significantly modified the results, reducing the raw frequency of "графіті" from 6,519 to 3,335 instances providing a clearer and more accurate representation of how "графіті" is used in Ukrainian epigraphic studies. The results of the methodological refinement highlight the importance of context-aware corpus analysis, particularly in specialised fields like epigraphy, where structural aspects of texts can heavily influence the interpretation of data.

We also refined the search of ***"напис"*** and ***"запис"*** to *напис* NOT ння* and *запис** respectively, to ensure that the search results excluded other words derived from the lemmas ***"напис"*** and ***"запис".***

5. Results.

5. 1. Comparative Analysis of the term "графіті", "напис", and "запис".

We conducted a detailed comparative analysis of the terms **"графіті", "запис", and "напис"**[1] using the Collocation frequency tool of the #LancsBox software package[2]. We determined the collocates in the network, more specifically the cloud of collocates for each term and the overlapping areas with the shared collocates.

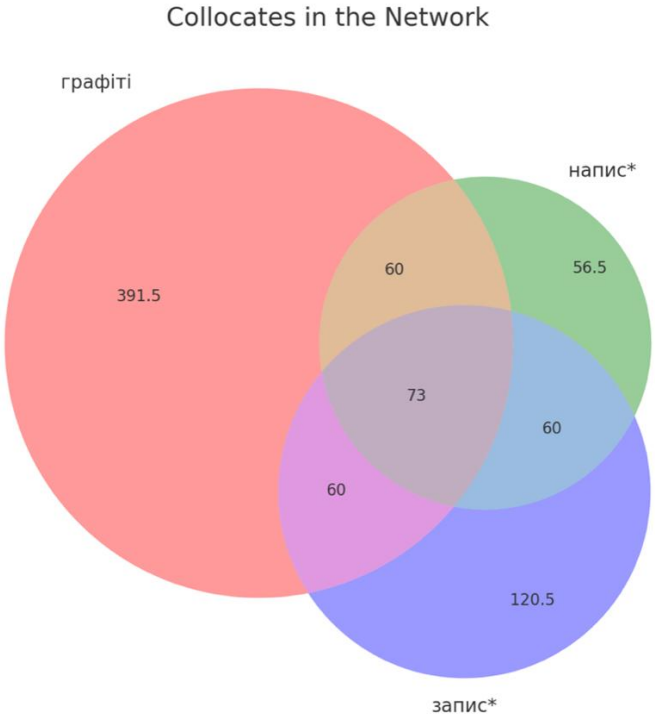


Figure 10: Collocates in the network.The overlapping areas represent the shared collocates:

1. Collocates shared by all three: 73 – *1, 2, 3, 4, 5, xiv, xv, бути, в, в.в, вибоями, виконання, встановити, від, він, графіті, датування, до, ж, з, за, його, категорія, київської, корнієнко, корпус, кінцем, кінця, лише, літер, літери, можемо, на, над, не, один, опис, оскільки,*

[1] The term "епіграфіка" was excluded from this analysis since its primary dictionary definition pertains specifically to the academic discipline.

[2] Brezina, V., Weill-Tessier, P., & McEnery, T. (2020). #LancsBox 5.x and 6.x [software]. Available at: http://corpora.lancs.ac.uk/lancsbox.
Brezina, V., Timperley, M., & McEnery, T. (2018). #LancsBox 4.x [software]. Available at: http://corpora.lancs.ac.uk/lancsbox.
Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics, 20*(2), 139-173.
Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge: Cambridge University Press.

2. Collocates shared by two nodes (divided equally between each pair): 120/2 = 60 - *xi, xvi, xvii, xvii, xii, xiv, xii, xiii, à, а, автор, автора, автором, аналогією, було, вертикальних, ви, виконаний, висоцький, виявлено, вказує, вочевидь, высоцкий, віднести, відносив, вірогідно, групи, датувати, двох, дві, для, дозволяє, другою, другої, же, засічки, засічок, зверху, знаходиться, комплекс, кутом, кількості, кінець, літера, лічби, мають, межах, методом, може, молитовний, між, місці, найвірогідніше, нами, напис, написання, них.*

[28]. The collocation network (*графіті NOT №, напис* NOT ння* and *запис** in Epigraphic corpus in Ukrainian (01 - Freq (5.0), L5-R5, C: 5.0-NC: 5.0)[1] and the analysis of that network has uncovered significant patterns and associations for these terms, shedding light on their usage in Ukrainian epigraphic discourse.

	Unique collocates	Shared among all	Shared with two
"графіті"	488	73	391,5
"напис*"	153	73	56,5
"запис*"	217	73	120,5

Figure 10: The relationships and linkages between these terms and their associated collocates

The analysis of unique collocates associated with the terms "графіті," "напис*," and "запис*" in the corpus reveals distinct patterns of term usage and association. 488 unique collocates are initially identified for the term "графіті". However, after accounting for 73 collocates that are shared among all three terms and dividing 120 collocates equally among the two other terms (indicating shared usage with either one of them), the net unique collocates for "графіті" stand at 391.5.

Following the same formula, "напис*" and "запис*" are left with a total of 56.5 and 120.5 unique collocates, respectively.

This calculation underscores the distinct lexical environments surrounding each term within the corpus, with "графіті" showing a significantly higher number of unique associations compared to "напис*" and "запис*." The substantial number of unique collocates for "графіті" suggests its broader or more varied contextual usage within Ukrainian epigraphic discourse. In contrast to "графіті", "напис*" and "запис*" exhibit fewer unique collocates, indicating a more focused or specialised use in the corpus.

Further analysis of the statistical correlations[2] (*The correlation analysis is based on individual files in the corpus (n = 17). p-values associated with the r (Figure 11) coefficients are available Figure 12*) between the collocates in the network points out a significant negative relationship between "графіті NOT №" and "напис* NOT ння," supporting the previous assumptions that these terms appear in different contexts or perform different text functions. There is also a minor negative correlation between "напис* NOT ння" and "запис*," reinforcing the idea of their distinct usage patterns.

[1] The full data is available in data/tsv/graphColl/sharedGraphColl_001.tsv.
[2] The data for the analysis is available in data/csv/statsAnalysis/multi_correlation.csv.

Variables	X.d0..b3..d1..80..d0..b0..d1..84..d1..96..d1..82..d1..96._NOT_.e2..84..96.	X.d0..bd..d0..b0..d0..bf..d0..b8..d1..81.._NOT_.d0..bd..d0..bd..d1..8f.	X.d0..b7..d0..b0..d0..bf..d0..b8..d1..81..
X.d0..b3..d1..80..d0..b0..d1..84..d1..96..d1..82..d1..96._NOT_.e2..84..96.	1.00	-0.68	-0.10
X.d0..bd..d0..b0..d0..bf..d0..b8..d1..81.._NOT_.d0..bd..d0..bd..d1..8f.	-0.68	1.00	-0.21
X.d0..b7..d0..b0..d0..bf..d0..b8..d1..81..	-0.10	-0.21	1.00

Figure 11: Pearson's correlations (r) between the search terms

Variables	X.d0..b3..d1..80..d0..b0..d1..84..d1..96..d1..82..d1..96._NOT_.e2..84..96.	X.d0..bd..d0..b0..d0..bf..d0..b8..d1..81.._NOT_.d0..bd..d0..bd..d1..8f.	X.d0..b7..d0..b0..d0..bf..d0..b8..d1..81..
X.d0..b3..d1..80..d0..b0..d1..84..d1..96..d1..82..d1..96._NOT_.e2..84..96.	-	0.0080	0.7236
X.d0..bd..d0..b0..d0..bf..d0..b8..d1..81.._NOT_.d0..bd..d0..bd..d1..8f.	0.0080	-	0.4740
X.d0..b7..d0..b0..d0..bf..d0..b8..d1..81..	0.7236	0.4740	-

Figure 12: P-values associated with Pearson's correlations (r)

5.2. Context-based analysis of the terms "епіграфіка", "графіті", "напис", and "запис".

As we established, all three terms share common collocates, including numerals and functional words, which show some overlap in their usage contexts. Yet, each term's unique collocates direct to different semantic fields or usage contexts. To fully understand the peculiarities of the usage patterns of these terms in academic discourse we conducted a detailed corpus-based analysis of their contextual application.

Analysis of the term "епіграфіка"

The first term of interest for our analysis is **"епіграфіка"**. In Ukrainian, the term "епіграфіка" (epigraphy) refers to an auxiliary historical and philological discipline that studies ancient inscriptions on hard materials, such as stone, metal and wood. The definition highlights epigraphy's focus on understanding and interpreting ancient writings and their significance in historical and linguistic research[1].

The term appears only 485 times in the corpus, with a relative frequency of 719.73. Its distribution is uneven, appearing only in 6 files out of the 17.

[1] 'Epigraphy', *Wikipedia*. Dec. 04, 2023. Accessed: Dec. 12, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Epigraphy&oldid=1188290541
'Епіграфіка', *Вікіпедія*. Aug. 14, 2023. Accessed: Dec. 12, 2023. [Online]. Available: https://uk.wikipedia.org/w/index.php?title=%D0%95%D0%BF%D1%96%D0%B3%D1%80%D0%B0%D1%84%D1%96%D0%BA%D0%B0&oldid=40154211

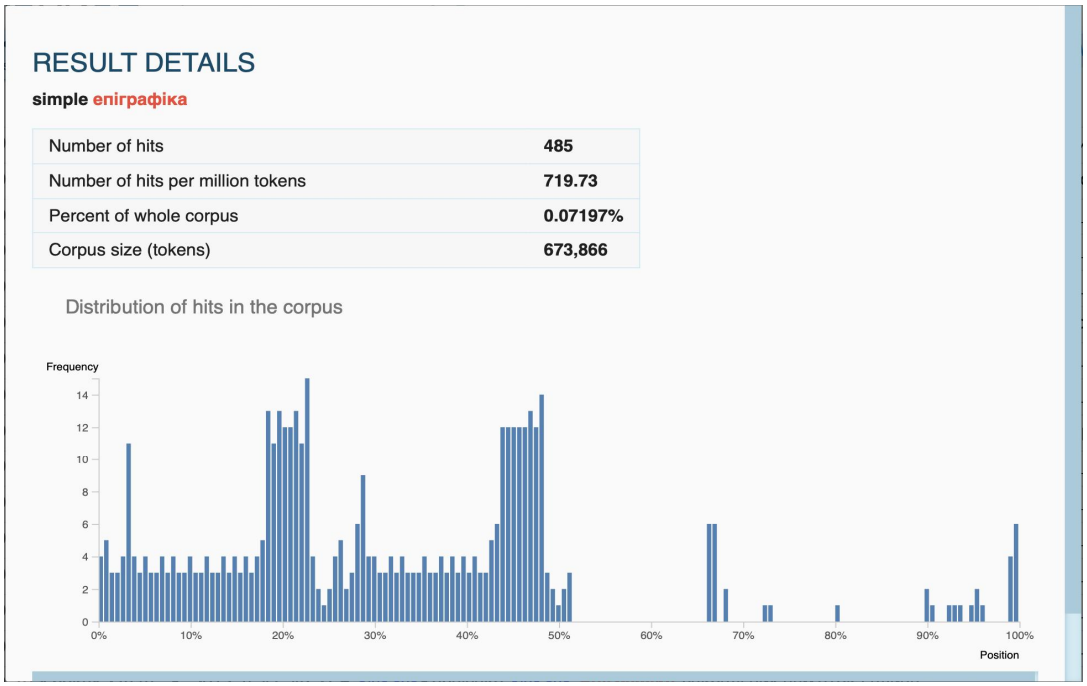


Figure 3: Statistical data for the term "епіграфіка"

Only in 3 instances out of 485 the term is used to refer to Epigraphy as a discipline.

*– **Епіграфіка (від грец. Επιγραφή – напис) вивчає** написи на твердому матеріалі. Всі такі написи давньоруського часу в –*

*– імена представників міщанських родин, які взагалі рідко фіксуються в письмових джерелах. Це посилює **роль епіграфіки для соціальної історії.** Очевидно, київськими міщанами були родини Шаїмових (графіті No 2062; текст: –*

What is more, 438 of these occurrences refer to the same bibliographic reference:

*– 419 Таблиця CCXXIII Фотографія графіті No 762 **В. Корнієнко. Епіграфіка сакральних пам'яток Галича –***

*– 421 Таблиця CCXXV Фотографія графіті No 763 **В. Корнієнко. Епіграфіка сакральних пам'яток Галича** –*

3 more instances refer to another bibliographic reference:

*– characteristics of their authors. Key words: biographical and genealogical studies, Saint Sophia, inscriptions. **ЕПІГРАФІКА Бібліографічне посилання: Скржинська М.В., Котляр М.Ф. ЕПІГРАФІКА [Електронний ресурс]** // Енциклопедія історії –*

The remaining samples the term is used for the body of inscriptions as a whole.

*– Шевченкове. Сучасний вигляд ПІСЛЯМОВА У книзі публікуються результати дослідження понад **восьми сотень пам'яток епіграфіки** XII –XIX ст. (графіті, дипінті, інскульпти), що збереглися на стінах споруд давнього Галича, таких як церква св. –*

*– збудовані у XVI ст. з кам'яних блоків, які походять з Успенського собору XII ст., тому на них **можна зустріти пам'ятки епіграфіки** ранішого часу, які потрапили до мурування разом з блоками розібраного храму. Загальна кількість пам'яток –*

– методологічною помилкою. Відтак, для уточнення часу появи тих чи інших форм написання літер ***в південноруський епіграфіці*** *ми залучаємо зібрані Я. Запаском описи рукописних книг XI–XVII ст., де автор наводить фотографії сторінок рукописів –*

– на тиньку прокреслена фігура у вигляді вертикальної заплетеної стрічки. Подібні малюнки часто ***зустрічаються в епіграфіці****, як приклад можемо навести два опубліковані графіті церкви Спаса на Нередиці [Овчаров 2006, с. 92, рис. 24–25].*

The analysis of the samples revealed that the term **"епіграфіка"** is employed in two different contexts. Surprisingly, the term is used only in three instances to refer to the primary dictionary meaning as an academic discipline. In the 483 cases the term refers to the body of inscriptions as a whole such as ***в південноруській епіграфіці, зустрічаються в епіграфіці, пам'ятки епіграфіки*** etc.

Analysis of the term "графіті"

The term in question **"графіті"** occurs 3,335 times in our corpus. It is evenly distributed across the corpus. This statistic and consistent occurrence across various sources are indicative of the term's widespread application and importance within the domain, and its pervasive presence in both academic literature and web-based discussions pertaining to Epigraphy within the Ukrainian context. This uniformity in distribution not only reinforces the term's centrality in epigraphic studies but also suggests its integral role in the discourse surrounding this field in Ukraine.

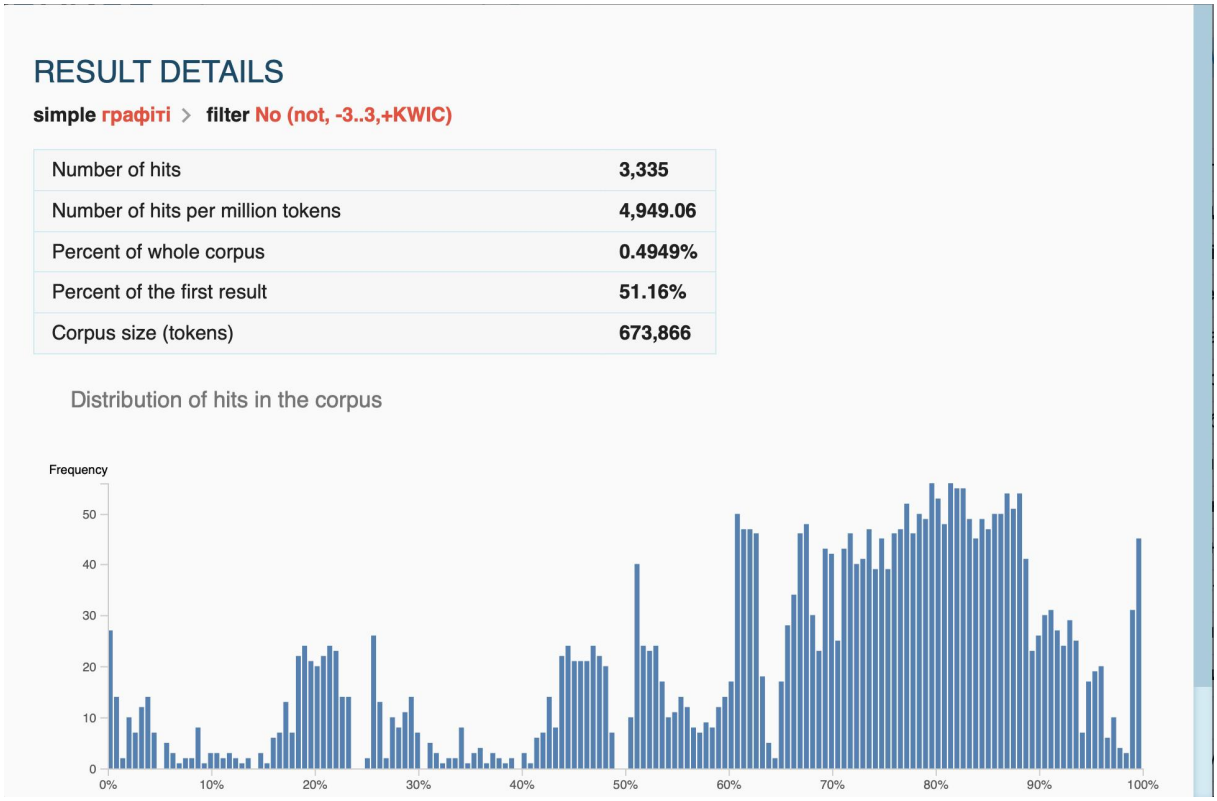
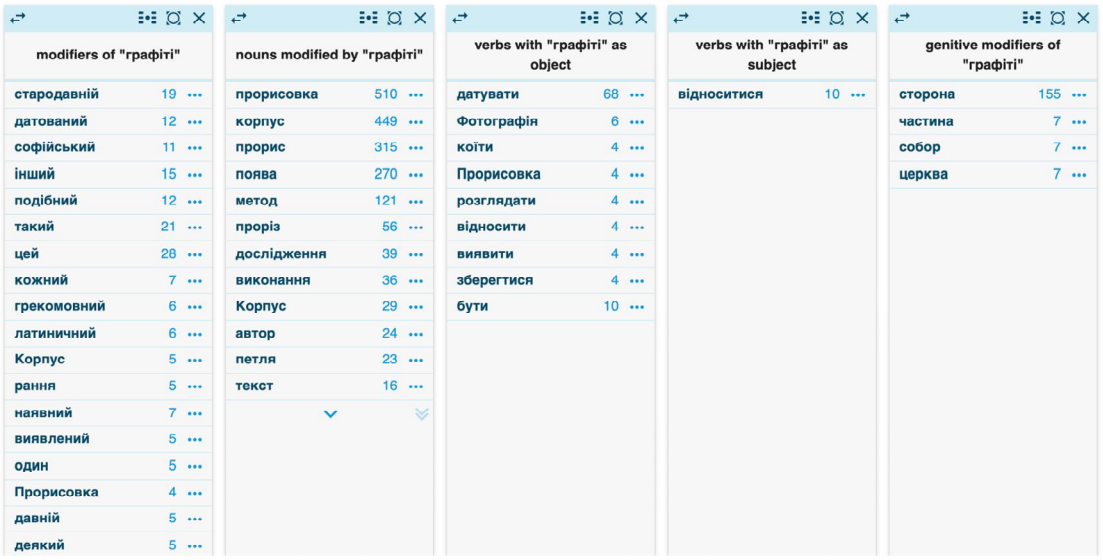


Figure 4: Statistical data for the term "графіті"



modifiers of "графіті"		nouns modified by "графіті"		verbs with "графіті" as object		verbs with "графіті" as subject		genitive modifiers of "графіті"	
стародавній	19	прорисовка	510	датувати	68	відноситися	10	сторона	155
датований	12	корпус	449	Фотографія	6			частина	7
софійський	11	прорис	315	коїти	4			собор	7
інший	15	поява	270	Прорисовка	4			церква	7
подібний	12	метод	121	розглядати	4				
такий	21	проріз	56	відносити	4				
цей	28	дослідження	39	виявити	4				
кожний	7	виконання	36	зберегтися	4				
грекомовний	6	Корпус	29	бути	10				
латиничний	6	автор	24						
Корпус	5	петля	23						
рання	5	текст	16						
наявний	7								
виявлений	5								
один	5								
Прорисовка	4								
давній	5								
деякий	5								

Figure 5: Collocations of the search term "графіті"

The analysis of the 10 most frequently occurring collocates with "графіті" suggests its use as a nominative term, especially in contexts involving

A. cataloguing or listing inscriptions or database entries of epigraphic findings;

– 127 Написи-графіті XVI – XVII століть центральної нави Софії Київської... Рис. 1. Фотографія та прорисовка ***графіті No 1884.*** *Пам'ятний напис Федора Ружинського. Перше-друге десятиліття XVII ст. Рис. 3. Прорисовка графіті No 1945. –*

B. classification methods

– No 2182 (табл. XLIII, 1) По фресковому тиньку прокреслені грецькі літери RLH. Датування. ***За типологічним методом графіті*** *датується в межах другої чверті XI – кінця XIII ст. Категорія: абетки, склади та окремі літери. Графіті No 2183 (табл. –*

C. the chronological and morphological analysis of inscriptions

– з них мають доволі широкий часовий діапазон поширення, який не впливає на обмеження ***хронологічних рамок появи графіті****. Для встановлення часу виконання графіті XVI – початку XVIII ст., окрім залучення епіграфічних матеріалів, –*

– {îzè}. Господи, поможи. Датування. За формою написання è з горизонтальною перемичкою, а також ***за структурою формули графіті*** *найвірогідніше датувати в межах другої чверті XI – кінця XV ст. Категорія: молитовний напис. Графіті No 2265 (табл. –*

D. a whole body of inscriptions

– цього остаточно пропонується авторський варіант реконструкції наявного на стіні тексту. 4 В. Корнієнко. ***Корпус графіті Софії Київської*** *–*

E. the location of the whole body of inscription

– Вказівник слів до ***графіті північної сторони*** *приділа святих Іоакима та Анни –*

So we can conclude that "графіті" in Ukrainian epigraphic discourse is not limited to a mere descriptive term of ancient writings but extends to a versatile term that plays an essential role in cataloguing, classifying, analysing and discussing inscriptions.

Analysis of the term "напис"

The term **"напис"** appears in the corpus 2,611 times, which translates to a relative frequency of 3,874.66. This substantial occurrence indicates its significant usage and

relevance in the epigraphic discours. The term is evenly distributed across the various documents. This pattern of occurrence highlights "напис" as a key term within the Ukrainian epigraphic discourse, used extensively in a majority of the texts, albeit not universally across the entire corpus. This level of frequency and distribution underscores the term's importance in academic and digital domains related to Epigraphy in the Ukrainian context, reflecting its integral role in the field.

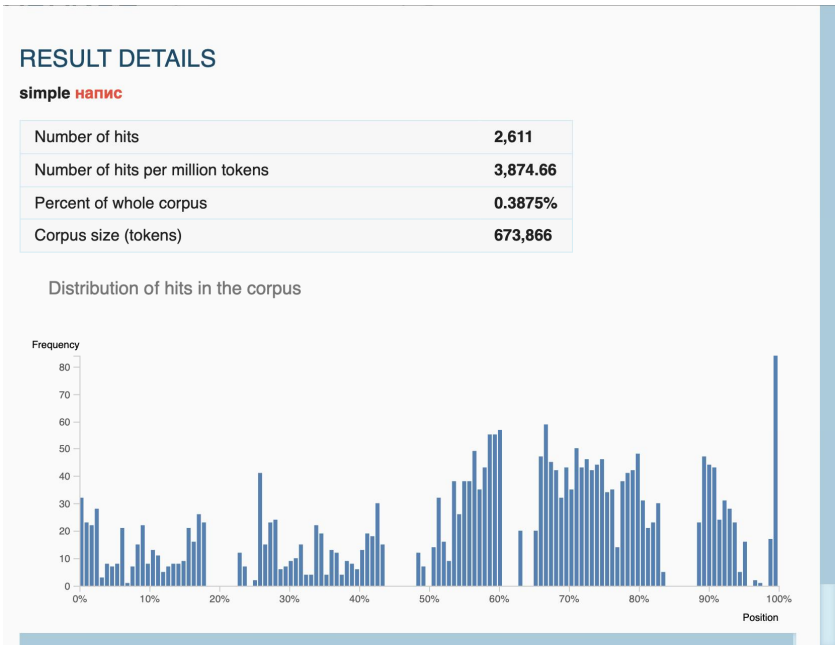


Figure 6: Statistical data for the term "напис"

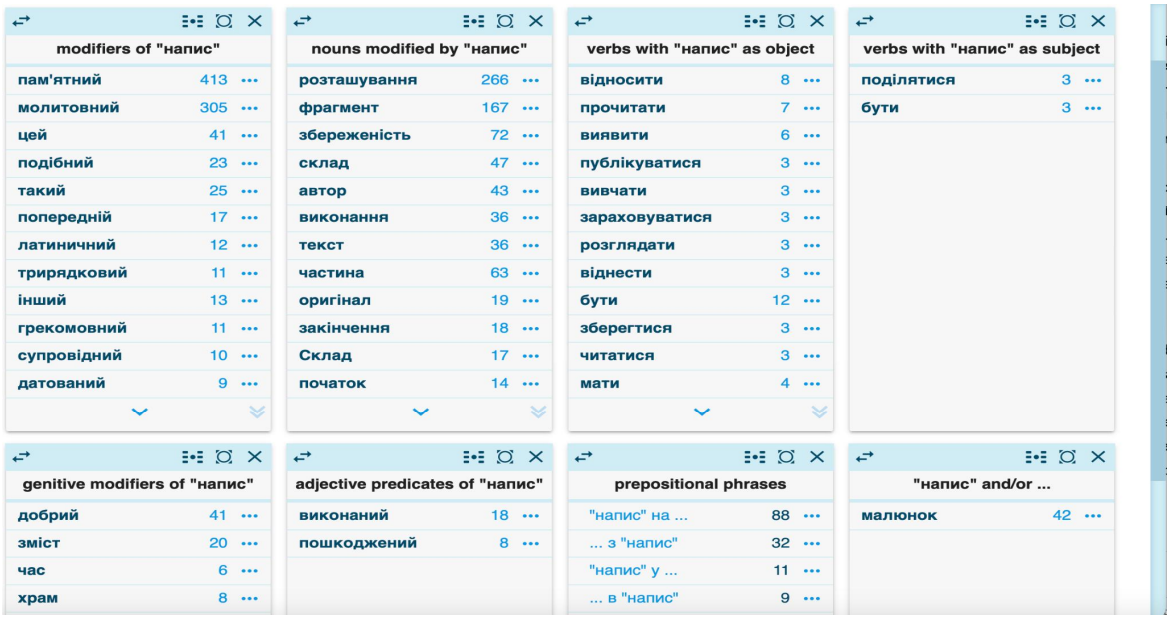


Figure 7: Collocates of the search term "напис"

The analysis of the first ten most frequently used collocations revels that the term "напис" is typically used for

A. classifying according to the type of the inscription: collocates like ***пам'ятні, молитовні, календарного змісту*** indicate its usage in academic context for classifying inscriptions according to content of the text of the inscription.

*– на фасадах храмів. Це, здебільшого, малюнки (зображення хрестів, тварин та емблем), окремі літери та **пам'ятні написи**, у текстах яких згадується низка представників відомих шляхетських родин Речі Посполитої XVI ст. Схожий у –*

*– за зовнішніми характеристиками графіті на менші категорії. Написи поділяються на такі категорії: 1) **молитовні написи**, які мають –*

*– 10) **написи календарного змісту**, що включають різні зразки розрахунку пересувних християнських свят, насамперед, Великодня; –*

B. describing the layout of the inscription: collocates like ***місці розташування, на стіні, на поверхні кам'яного блоку, в місці розташування*** point towards its use in the contexts describing the location of the inscription on the monument.

*– на висоті 145 см від рівня ґрунту та на відстані 145 см від правого кута. Опис. Поверхня стіни **в місці розташування напису** сильно пошкоджена, частина складових не простежується: ѱ҃авъкѣ҃вавъг – З розбивкою на слова та відновленням –*

*– султана. Оскільки найзручніший шлях до Порти пролягав через Галич, вочевидь, в один з таких візитів і з'явився **написи на стіні** костелу св. Станіслава. Можливо, він вплинув на короля в питанні передачі занедбаного на той час костелу –*

C. describing the structure of the inscription: collocates like ***фрагменти, частина, початок*** indicate that the term is used in the contexts describing the structure of the inscription.

*– 13) **фрагменти написів**, куди відносяться пошкоджені графіті, від яких збереглося декілька літер або окремих слів, що не дозволяють –*

*– 141 Починаючи з літер ги ця **частина напису** не була прочитана С. Висоцьким тв. В. Німчуком. 142 С. Висоцький та В. Німчук читали тут ь. 143 Замість літер сво С. –*

D. describing the state of the preservation: the collocations like ***збереженість*** indicate that this term is used also while describing the state of preservation of the inscription.

*– **Збереженість напису** доволі добра, його складові чітко читаються: илипъ Коровинъ. Пилип Коровін. Праворуч помітна велика літера Л, –*

These varied uses of "напис" mentioned above, though not exhaustive, demonstrate that the term is used for describing, classifying, relating separate items of epigraphic inscriptions to their contents, function or layout. Moreover, the distribution of "напис" across the corpus reflects its multifaceted role in academic discussions, documentation, and analysis of epigraphic materials.

The study's collocation network illustrates the relationships between "напис" and its associated terms, mapping out a complex web of connections that underline its importance in scholarly research. This reinforces the term's central position in the academic study and communication regarding epigraphic materials, especially in historical and analytical frameworks.

Analysis of the term "запис"

The term **"запис"** stands out within the Ukrainian Epigraphic corpus, with its occurrence 1,079 times with relevant frequency 1,601.21 in 13 out of 17 texts. However, the term is not evenly distributed with a significant concentration of occurrences in the files related to V. V. Korniienko's series on the epigraphic heritage of Ukraine (Korniienko, 2010b).

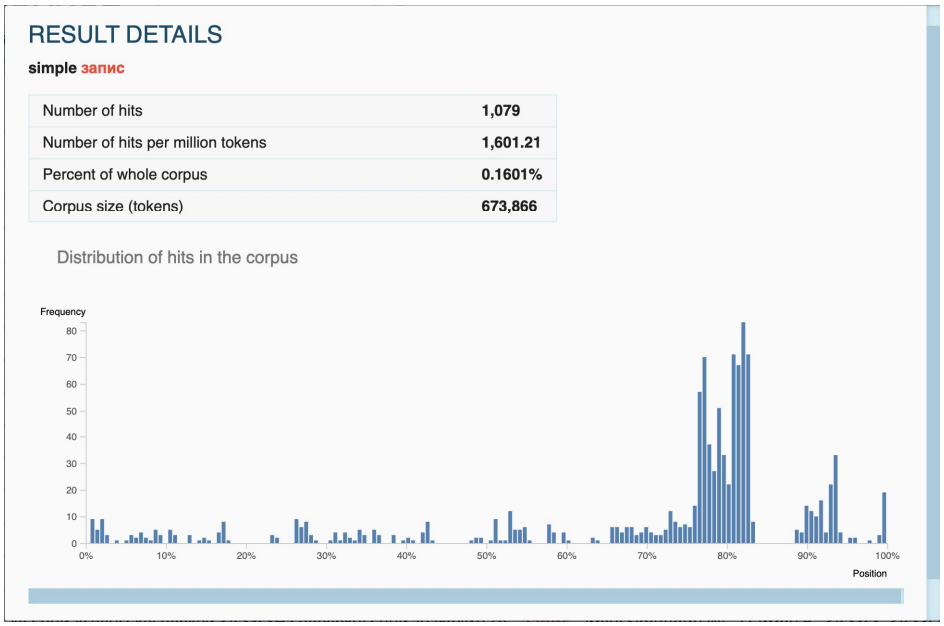


Figure 8: Statistical data for the term "запис"

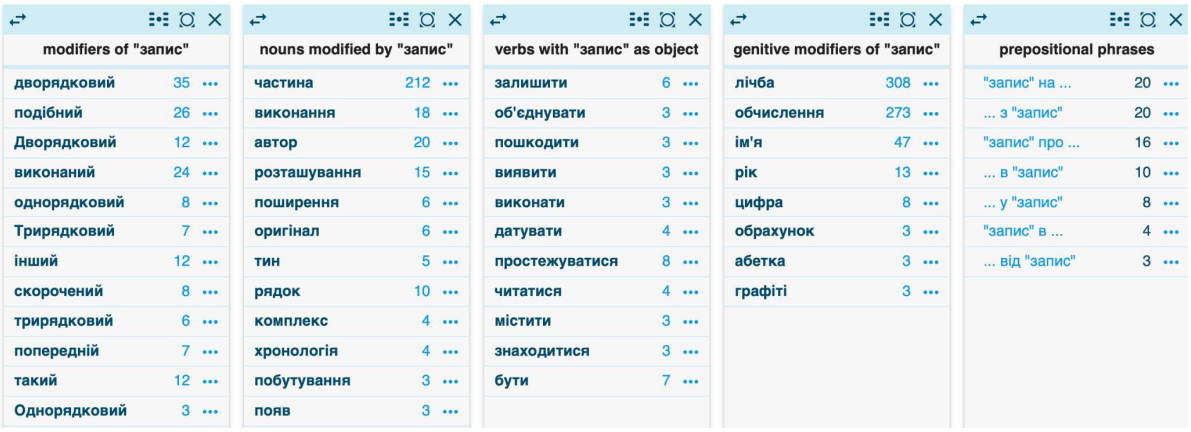


Figure 9: Collocations of the search term "запис"

The analysis of the 10 most frequent collocations points to the term's role in

A. <u>in organising numerical data and engaging in classification:</u>

This is evident from collocations with terms such as ***"лічба", "обчислення", "кількості", "дворядковий".***

– слід також датувати кінцем XI – серединою XII ст. Категорія: ***запис лічби.*** *Комплекс розташованих один над одним записів лічби, що містить чотири графіті (NoNo 2651–2654). –*

– Опис. На поверхні стіни у стовпчик прокреслені тридцять вісім горизонтальних засічок ***рахункової частини запису обчислення кількості*** *поминальних служб. Датування. Кінець XII – кінець XIV ст.; за хронологією поширення подібних –*

The examples illustrate the categorization of a series of counts and descriptions of inscriptions containing numerical data for calculating quantities or memorial services,

including a specific focus on the numerical and quantitative aspects within the epigraphic content.

B. referring to the authors' of the inscriptions:

*– малюнок хреста, що міг бути розташований на верхньому блоці вівтарної перегодки. В. Вуйцик вважав **автором запису єпископа Козьму**, поставленого Галичу 1157 р., про що повідомляє літопис: того же лѣта поставленъ бысть епископъ –*

The term is mentioned in the contexts related to the authors of the inscriptions, as in the example by referencing historical figures believed to be the author of specific inscription. This use of **"запис"** highlights its application in attributing authorship and historical context to inscriptions.

This usage of the term extends beyond describing the epigraphic content, playing an essential role in academic discourse where the citation of the sources and the referencing of specific pages and entries are important for substantiating research findings and discussions.

Further analysis of the corpus samples reveals that the term **"запис"** is predominantly used in contexts referring to the inscriptions associated with numerical data, numbers, counting, personal names and pages in bibliographic references. This suggests that the term often denotes the text of the inscription as a note, which contains information about dates, people, and various counts. Moreover, it is also involved in the academic discourse surrounding epigraphy, where citing sources and referencing specific pages or entries is crucial. Such multifaceted usage of the term "запис" indicates its significance not only in the direct discussion of the epigraphic content, but also as an essential element in academic discourse surrounding epigraphy.

The term's connections with "графіті" and "категорія" underline its integral role in the descriptive and classificatory processes, serving a dual purpose of methodological documentation and scholarly citation within Ukrainian epigraphy literature.

5.3. SKOS Concept Schema: UkrainianEpigraphicVocabulry.

Based on the initial findings from our investigation, we have modelled the principal four entries of the SKOS taxonomy for epigraphy (*D4.2_Guidelines for Mapping into SKOS.Pdf*, n.d.), which encapsulate the core concepts and their interrelations within the field:

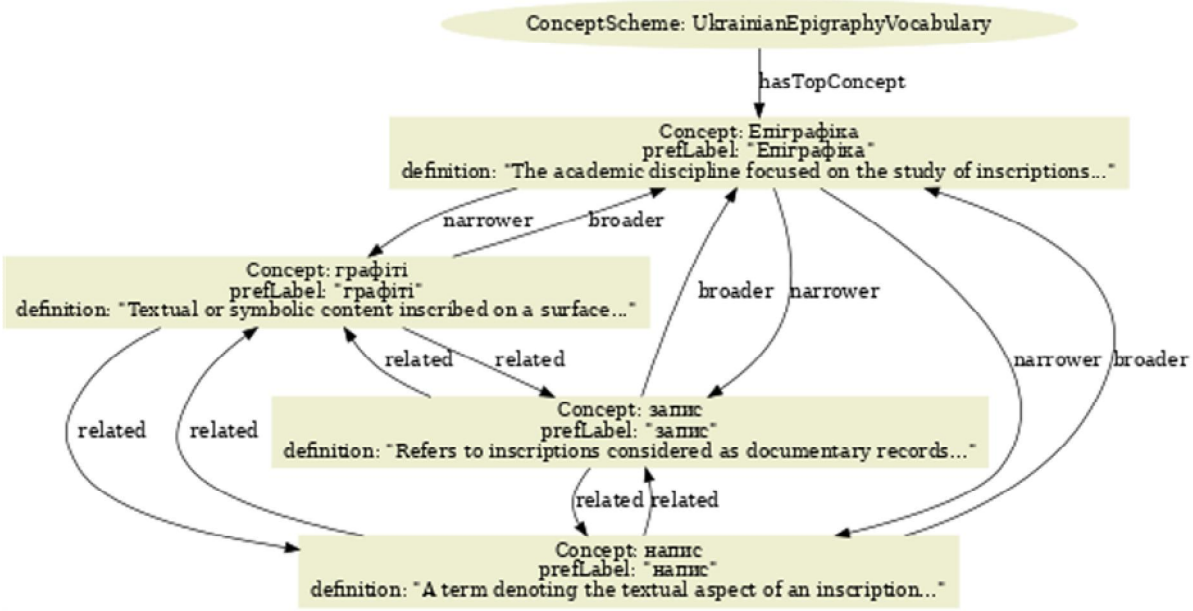


Figure 10: ConceptScheme: UkrainianEpigraphyVocabulary

At the top, we have the "ConceptScheme: UkrainianEpigraphyVocabulary", which serves as the overarching framework for the epigraphic terminology. From here, there are "hasTopConcept" connections to the primary concept in the scheme:

- **Concept: Епіграфіка (Epigraphy)**
  prefLabel: "епіграфіка" – This is a preferred label for the concept, indicating the canonical name used in the dataset.
  definition: "The academic discipline focused on the study of inscriptions..." – This provides a brief explanation of what the field of epigraphy encompasses.

Branching from the central concept of epigraphy, we have three related concepts1:

1. **Concept: графіті (Graffiti)**
   a. **prefLabel: "графіті"**
   b. **definition:** "Textual or symbolic content inscribed on a surface..." – This term is used for actual inscriptions or graffiti, the primary subject of epigraphic studies. It's connected to epigraphy as a narrower term, indicating that graffiti is a subset of the broader field of epigraphy.
2. **Concept: запис (Record)**
   a. **prefLabel: "запис"**
   b. **definition:** "Refers to inscriptions considered as documentary records..." – This suggests that the term is used for inscriptions that serve a documentary purpose, perhaps recording events, names, or data. This is related to both the broader field of epigraphy and to other specific terms.
3. **Concept: напис (Inscription)**
   a. **prefLabel: "напис"**
   b. **definition:** "A term denoting the textual aspect of an inscription..." – This term likely refers to the content aspect of inscriptions, emphasising the written message itself.

The relationships between these terms are depicted as bidirectional arrows, indicating that the understanding of one term may influence the interpretation of others. The use of "broader" and "narrower" suggests a hierarchical relationship where one term is a subset of the other, while "related" signifies a non-hierarchical associative link.

This SKOS concept schema illustrates the intricate structure of terms within the Ukrainian epigraphic domain, showing how various terms relate to one another in the context of epigraphic research. It emphasises the importance of clear definitions and relationships to ensure accurate and consistent use of terminology in the field.

6. Discussion.

The results of our study demonstrate the difficulties associated with establishing a standardised vocabulary, mostly due to language variations and the influence of individual scholars' "personl vocabularies", particularly within the field of epigraphy. This makeit essencial to adapt a suitable methodology that takes into account the linguistic, historical, and cultural background. The leverage of digital technologies and corpus-based analysis has demonstrated significant value in offering relevant insights that can inform the creation of harmonised / standardised vocabularies and contribute to the wider discourse surrounding the digitizaiton of epigraphic heritage.

[1] We have not included the Alt:Labels, detailed definitions and documentaion in this sample SKOS schema, which serves solely as a test of our methodology.

The analysis of the synonymous terms , **“графіті”, “запис” and “напис”** as well as the term **“епіграфіка”** within our sample corpus was essential for establishing both similarities and differences in their usage. The integration of the quantitative and qualitative methodologies has enhanced our understanding of the hierarchical structure and interrelations among these terms/concepts, hence playing a crucial role in the development of the simple model for SKOS.

The results of the analysis demonstrate the importance of the context when examining the usage patterns of terminologies, since each term carries its own unique implications despite being part of a synonymous set. The term **“графіті”** is frequently associated with the tangible part of the inscription, which are characterised by their historical and archeological significance. “Напис” typically accentuates the textual composition and structure of the inscription, whereas “запис” is often associated with the role of inscriptions as documented records, incorporating numerical information or historical records.

The field of **“епіграфіка”** provides a comprehensive framework for studying these inscriptions, offering a broader academic viewpoint on the specialised terminologies used in epigraphy. A specific minor segment of the SKOS model has been designed to replicate these subtle associations, offering a semantic framework that encompasses the complexities of epigraphic terminology. This preliminary research and the corresponding segment of the SKOS model not only encompass the lexical subtleties found in our corpus but also showcase the effectiveness of our methodology in establishing the foundation for semantic precision and integration within the domains of digital humanities and the preservation of cultural heritage.

7. Conclusions.

Based on the discussions and insights derived from the comprehensive analysis of the synonymous triplet **“графіті”, “напис”, and “запис”** within Ukrainian epigraphy, this study concludes with several key observations and implications for the field of digital epigraphy, digital humanities and the standardisation of terminologies.

First, the nuanced use of **“графіті”, “напис”, and “запис”** underscores the critical role of descriptive terminology in cataloguing, classifying, analysing, and discussing epigraphic inscriptions. Each term, through its specific usage and collocates, highlights distinct facets of epigraphic work, ranging from the documentation of ancient writings to the organisation of numerical data and the academic discourse surrounding these inscriptions.

Second, the study reaffirms the profound impact of linguistic variations and cultural contexts on using and understanding epigraphic terminology. It emphasises the importance of considering these factors in developing a harmonised/standardised vocabulary that can accommodate the richness and diversity of epigraphic traditions across different regions and languages.

Further, digital methodologies, including corpus-based analysis, have been instrumental in uncovering the intricate patterns of term usage within Ukrainian epigraphy. These tools not only enhance the scope and accuracy of research but also pave the way for more nuanced and comprehensive studies in the digital era.

Implications for Future Research and Practice:

This study contributes to the ongoing efforts to standardise epigraphic terminology, offering a foundation for developing controlled vocabularies that balance linguistic diversity with the need for clarity and consistency in academic discourse.

Moreover, by highlighting the linguistic and cultural nuances of key epigraphic terms, the research fosters a deeper appreciation for the complexity of epigraphic studies. It

encourages cross-cultural collaboration and dialogue, essential for enriching the global understanding of epigraphy.

The findings underscore the potential of digital tools and methodologies to transform epigraphic research. Future studies are encouraged to leverage these technologies further to explore uncharted territories in epigraphy and to facilitate the preservation and interpretation of cultural heritage in the digital age.

Moreover, the finel results of our research will help to establish a benchmark for SKOS creation for other domanins, specialy involved with digital cultural heritage and digitization humanities in general.

References

A New Companion to Digital Humanities (1st ed.). (2015). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118680605

Bates, M. J., & Maack, M. N. (Eds.). (2009). Controlled Vocabularies for Art, Architecture, and Material Culture. In *Encyclopedia of Library and Information Sciences, Third Edition* (0 ed., pp. 1277–1281). CRC Press. https://doi.org/10.1081/E-ELIS3-120044074

D4.2_Guidelines for mapping into SKOS.pdf. (n.d.). Retrieved November 20, 2023, from https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/ATHENA/Deliverables/D4.2_Guidelines%20for%20mapping%20into%20SKOS.pdf

Dobreva, M. (2005). *Digitisation of Cultural Heritage: Between EU Priorities and Bulgarian Realities*. https://www.academia.edu/2855207/Digitisation_of_Cultural_Heritage_between_EU_Priorities_and_Bulgarian_Realities

Elliott, T. (2014). Epigraphy and Digital Resources. In C. Bruun & J. Edmondson (Eds.), *The Oxford Handbook of Roman Epigraphy* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195336467.013.005

European Commission report on cultural heritage: Digitisation, Online Accessibility and Digital Preservation. (2019, June 27). European Heritage Tribune. https://heritagetribune.eu/europe/european-commission-report-on-cultural-heritage-digitisation-online-accessibility-and-digital-preservation/

Heřmánková, P., Horster, M., & Prag, J. (2022). *Digital Epigraphy in 2022: A Report from the Scoping Survey of the FAIR Epigraphy Project*. Zenodo. https://doi.org/10.5281/zenodo.6610696

Korniienko, V. V. (2010a). *Korpus hrafiti Sofiï Kyïvs′koï (XI - pochatok XVIII st.) = Collection of graffiti of St. Sophia of Kyiv*. Natsional′nyĭ zapovidnyk "Sofiia Kyïvs′ka" ; Instytut ukraïns′koï arkheohrafiï ta dzhereloznavstva im. M.S. Hrushevs′koho.

Korniienko, V. V. (2010b). *Korpus hrafiti Sofiï Kyïvs′koï (XI - pochatok XVIII st.) = Collection of graffiti of St. Sophia of Kyiv*. Natsional′nyĭ zapovidnyk "Sofiia Kyïvs′ka" ; Instytut ukraïns′koï arkheohrafiï ta dzhereloznavstva im. M.S. Hrushevs′koho.

Lerones, P. M., & Llamas, J. (2021). Digitizing cultural heritage: What are we talking about? *Academia Letters*. https://www.academia.edu/50926326/Digitizing_cultural_heritage_What_are_we_talking_about

Marinković, B., & Maslikovic, D. M. (2019). Cultural Heritage Digitization in Serbia: Standards, Policies, and Case Studies. *Sustainability*. https://www.academia.edu/93587175/Cultural_Heritage_Digitization_in_Serbia_Standards_Policies_and_Case_Studies

Miric, A. (2016). EUROPEAN POLICY ON DIGITISATION OF CULTURAL HERITAGE FROM 2005. ONWARDS. *FACTA UNIVERSITATIS Series: Law and Politics*. https://www.academia.edu/35101737/EUROPEAN_POLICY_ON_DIGITISATION_OF_CULTURAL_HERITAGE_FROM_2005_ONWARDS

Mugridge, R. L. (2007). Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images. *Library Collections, Acquisitions, & Technical Services*, 31(3–4), 230. https://doi.org/10.1080/14649055.2007.10766172

Soriano, I. V., & Espinosa, D. E. (Eds.). (2021). *Epigraphy in the Digital Age: Opportunities and Challenges in the Recording, Analysis and Dissemination of Inscriptions*. Archaeopress Publishing Ltd. https://doi.org/10.2307/j.ctv1xsm8s5

ЕПІГРАФІКА. (n.d.). Retrieved November 29, 2023, from http://resource.history.org.ua/cgi-bin/eiu/history.exe?&I21DBN=EIU&P21DBN=EIU&S21STN=1&S21REF=10&S21FMT=eiu_all&C21COM=S&S21CNR=20&S21P01=0&S21P02=0&S21P03=TRN=&S21COLORTERMS=0&S21STR=Epigrafika

Бібліографічний опис:

Тамразян, А. С. (2024). Описова термінологія в українській епіграфіці: корпусне дослідження синонімічного триплету “графіті”, “напис” і “запис”. *Науковий часопис Українського державного університету імені Михайла Драгоманова. Серія 9. Сучасні тенденції розвитку мов*, 27, 92–110. https://doi.org/10.31392/NPU-nc.series9.2024.27.08.

Анотація

Основною метою дослідження є аналіз описової термінології, яка використовується в академічному дискурсі, присвяченому українській епіграфіці, із застосуванням корпусного підходу. Використовуючи інструмент BootCat і користувальницькі кортежі, такі як “епіграфіка графіті напис” і “корпус Україна графіті”, ми зібрали веб-сайт і згенерували зразок корпусу, що містить українські тексти, пов’язані з епіграфікою. Мета статті полягає у вияві епіграфічних термінів, що циркулюють у цифровій сфері, а також аналізі особливостей їхнього використання. Виявлено варіації у вживанні термінів в українських епіграфічних текстах, запропоновано бачення їхнього академічного застосування. У дослідженні було продемонстровано, як синоніми можуть не лише збагатити мову, але й створювати проблеми у стандартизації термінології через лінгвістичні варіації та вплив окремих учених. Результати підкреслюють необхідність нюансованого підходу до розробки стандартизованих словників з урахуванням мовного, історичного та культурного контекстів. Це дослідження є цінним прикладом того, як цифрові інструменти можуть допомогти зрозуміти та вирішити труднощі, властиві збереженню та вивченню культурної спадщини в епоху цифрових технологій.

***Ключові слова**: гармонізована/стандартизована термінологія, SKOS, цифрова епіграфіка, корпусна лінгвістика.*