The Doorway to Deciphering: Methodology and Analysis of Successful Decoding Results in the Voynich Manuscript.

Alisa Gladyševa *

*Faculty of Philology, Vilnius University, Vilnius, Lithuania.
https://orcid.org/0000-0001-7739-1714
DOI: https://doi.org/10.55248/gengpi.5.0624.1511

ABSTRACT

Despite all the unsuccessful decipherment research methods that others have tried to apply, there was no solution in reading the Voynich manuscript. This research describes recent successful decoding methodology from the very doorway to deciphering the whole Voynich manuscript. In particular, the investigation of background information and detailed analytical studies of the characters of the Voynich manuscript, words that were written in plain way, which have been performed to define the algorithm of the encoded text. It gives emphasis on the importance of this research for its successful decipherment.

Keywords: Voynich manuscript, MS 408, text decipherment, codicology, cipher.

1. Introduction

Given the research embargo period on access to the book titled The deciphering the whole text of the Voynich manuscript by Alisa Gladyševa: the linguistic basis of the text, index, and analysis, isbn: 978-609-475-419-7, this research article represents one of the initial contributions in a series of open-access publications by the author, focusing on the Voynich manuscript, which she has already deciphered herself.

The Voynich Manuscript, known as Yale University’s Beinecke Library Manuscript 408, remained undeciphered for centuries [1]. This research describes recent successful decoding approaches from the very doorway to deciphering the whole Voynich manuscript. Decipherment of the Voynich Manuscript is widely considered to be the most important achievement of the new innovative research.

The Voynich Manuscript is an illustrated manuscript. The manuscript itself is bound in vellum, it was rebounded lately with the obvious loss of some folios. For this moment, it contains 116 folios, it means 232 pages. It was found by Wilfrid Voynich of Lithuanian origin, at the villa housed a former Jesuit library Villa Mondragone in 1911 or 1912. In 1969, the manuscript which bears Voynich name was donated to Yale’s Beinecke Library by H. P. Kraus.

According to the results of innovative research of the Voynich manuscript was written in medieval Galician (Galician-Portuguese) (Gladyseva, 2019). The most important result in the encryption code of the Voynich manuscript that made decryption more complicated over the centuries was the use of specific cipher algorithm of polyalphabetic encryption that was used partially and simultaneously with monoalphabetic encryption (Gladyseva, 2020). The implementation of this specific cipher algorithm by the master of the Voynich manuscript resulted in the transformation of numerous words into plaintext, implying that they were effectively left unencrypted.

The research results generating considerable interest in the very decipherment methodology itself. In particular, the doorway to investigation of background information and detailed analytical studies of the characters ‘l’, ‘a’, ‘v’, ‘n’, ‘d’ of the Voynich manuscript, their decipherment, the reading and examination of vernacular plant names within the text, which contain these symbols. The results of analysis demonstrate that as a consequence of the particular cipher algorithm employed by the master, occasionally led these symbols to their conversion into plaintext characters. Introduction should be typed in Times New with font size 10. In this section highlight the importance of topic, making general statements about the topic and Presenting an overview on current research on the subject. The simplest way is to replace(copy-paste) the content with your own material. Your introduction should clearly identify the subject area of interest.

2. Methodology

Some researchers focus solely on determining the language corresponding to the Voynich manuscript, yielding results that only rightly confirm the manuscript’s human origin (R. Carbo-Dorca, 2021: 5). Robert Downing et al. (2020) accurately asserted in their article that the Voynich manuscript is composed in an Indo-European language (Downing et al., 2020). Others, such as Rugg and Taylor (2016), Daruka (2020), Timm and Shinner (2020)
have not achieved conclusive results, arguing that there might not be any discernible language in the Voynich manuscript. Gibbs (2017) proposes that the manuscript is a form of abbreviated Latin.

The initial attempts at deciphering date back to the 16th century, but large-scale studies have consistently failed due to challenges in determining the structure of the text and the language in which it was written. Hence, emphasizing the significance of the results obtained in this study as a notable decipherment of the Voynich manuscript becomes crucial.

Within the domain of academic inquiry, in their article, The Linguistics of the Voynich Manuscript, Claire L. Bowern and Luke Lindemann from the Department of Linguistics at Yale University label the Voynich Manuscript as a curious topic for linguistics research, given its absolutely unknown language (Bowern, & Lindemann, 2021: 2). Notably, they assert that word and line-level metrics indicate a regular natural language, marking a substantial breakthrough in Voynich Manuscript scientific research (Bowern, & Lindemann, 2021: 17).

However, the assumption that a line of text roughly corresponds to a sentence is incorrect, as the Voynich Manuscript text is written line-by-line, not approximately equating to a sentence. Therefore, the assertion of agreement appearing more frequently in the last word of the line is inaccurate, as one line does not necessarily represent one sentence, as commonly believed (Bowern & Lindemann, 2021: 17).

The researchers also suggest that blocks of running text are separated into paragraphs, potentially indicating a shift in topic (Bowern & Lindemann, 2021: 16). Contrary to this, innovative research on decipherment suggests that the paragraphs on botanical sheets represent different facets of the same topic, showcasing rare nuances. Some folios feature descriptions of several similar plants in consecutive paragraphs, each detailing the characteristics of different plants but of the same genus, as it will be possible to see further.

The researcher who came closest to deciphering the entire Voynich Manuscript was Dr. Erwin Panofsky. Initially, he proposed the idea that the manuscript is written in Spanish, potentially originating from the southern regions near Spain, and suggested an influence from Jewish Kabbala. In his 1932 comments on the names of the months, he noted: “...which certainly suggests some form of Spanish, rather than Latin or French...” (Panofsky, 1932). However, in Military cryptanalysts: Panofsky's responses of 1954, Panofsky changed his stance and erroneously claimed that the names of the months were written in 'provincial French' (Panofsky, 1954).

According to Gibbs (2017), the Voynich manuscript is a form of abbreviated Latin.

In the landscape of linguistic inquiry, a dominant hypothesis emerges in 2022 as Yale University researchers Daniel Gaskell and Claire Bowern present experimental findings. Their study involves participants intentionally generating nonsensical text, revealing patterns mirroring statistical properties ostensibly "found" in the Voynich manuscript, hinting at potential hoax origins (Bowern & Gaskell, 2022).

Diverging from the limited advancements made by researchers who struggled with deciphering, the proposition that the language of the Voynich Manuscript aligns with the medieval Romance family garners substantial attention. A significant decipherment in this discourse was made in 2019, "Deciphering the Whole Text of the Voynich Manuscript" (Gladyseva, 2019), with proposing a medieval Galician (Galician-Portuguese) origin with the usage of Vulgar Latin, Old Spanish, Old French, and Old Catalan plant nomenclature within the manuscript. This exhaustive study includes transliterations, translations, an intricate linguistic analysis, and a compelling proposal, positing the manuscript as a late Medieval manual designed for alchemical purposes (Gladyseva, 2019: 11).

In the pursuit of deciphering Voynich manuscript texts, an article by Gerard Edward Cheshire introduces a theory wherein he asserts, "The version of the proto-Romance language used in the MS408 manuscript..." (Cheshire, 2019, p. 36). However, Cheshire faced substantial criticism from notable scholars, such as Dr. Lisa Fagin Davies, who dismissed his Proto-Romance theory as nonsensical. Additionally, Alisa Gladyseva, in her article (2020), contradicts Cheshire's claim by asserting that the Voynich manuscript was written in medieval Galician: "The Voynich manuscript was not written in the proto-Romance at all, but in a language that belongs to the Western Romance subdivision, particularly to the Iberian-Romance branch, Western group: the medieval Galician language - Galician-Portuguese" (Gladyseva, 2020).

In a notable development in November 2021, significant events unfolded, prompting careful consideration within scholarly discourse. Cheshire's apparent endorsement of Alisa Gladyseva's research on the medieval Galician origins of the Voynich manuscript's language raises concerns of potential "borrowing", as he published a study stating: "It has transpired since, that the language is more precisely a hybrid of an Iberian variant of Vulgar Latin, known as Galician-Portuguese" (Cheshire, 2021: 2) without any reference to Gladyseva's much earlier works, cf. A. Gladyseva (2018; 2019).

3. Analysis

Methodological analysis upon the problem of deciphering the Voynich Manuscript for over 500 years. Outlining the past history of the study, previous solutions to decipherment of the Voynich Manuscript are inadequate, incorrect, and oversimplistic.

In this chapter is analyzing the most basic description of the problem phenomena of deciphering the Voynich Manuscript for over hundreds years. Highlighting limitations of previous studies. The main cryptographic problems:

(1) The main researchers' problem was undefined language. It was unknown what language it was written, in which language branch the language of Voynich manuscript belonged to. Even in a new researches for eg. Yale University research by C. L. Bowern & L. Lindemann, (2021) it was not even known whether it belongs to the Indo-European linguistic family tree or not.
(2) The paleography of the invented letters’ characters of the Voynich Manuscript is unknown due to inexistence closely similar medieval manuscript.

(3) Difficulty in decryption due to symbols’ similarity, and as a result wrong alphabet paleography determination. E.g. various transcription alphabets have been created to equate it characters with Latin characters such as the European Voynich Alphabet (EVA).

(4) Impossibility of making it machine readable due to unknown algorithm rules for the paleography determination of the characters. E.g. The “First Study Group” led by cryptographer William F. Friedman in the 1940s attempted to transcribe each line of the Voynich manuscript to an IBM punch card to be machine readable. The failure to make the Voynich handwritten machine readable has several issues that will be mentioned below.

(5) The terminology. A lot of technical-botanical terms, that were written in medieval Galician language.

(6) Frequent use of the nomenclature of vernacular plant names in medieval Galician, but also using Spanish, Aragonese, Catalan, Euskara vernacular names of plants instead of the Latin nomenclature of plant names, which was better known in medieval times from the ancient times of Dioscorides, Strabo, Theophrastus and etc. Material which are used is presented in this section. Table and model should be in prescribed format.

4. Results and discussion: the doorway to successful decoding results and the very methodology of Voynich Manuscript decipherment.

The approach adopted in this study utilized a method that allowed to resolve the cipher. Gianbattista della Porta (1535? – 1615) posited that when the subject matter of the coded text is identified, analysts can make informed hypotheses regarding the words commonly employed within a specific technical context. The connotation of each text is delineated by recurring terms that hold significance within the given discourse.

These deciphering methodology assumed particular significance in the analysis of the Voynich manuscript, particularly in light of the botanical references undoubtedly present in its mostly initial sections.

The paramount methodology essential for deciphering the Voynich code centered on the interpretation of plant vernacular names, alongside the identification of vernacular plant names that could potentially have been written in plaintext, due to the cipher algorithm devised by the master of the Voynich manuscript. Notably, during the initial phase of deciphering the entire Voynich Manuscript, the cipher algorithm remained entirely unknown and was only revealed subsequent to the determination of certain plaintext paleography through characters' analysis in this scholarly research.

Fascinatingly, Porta invented methodology of breaking the multi-alphanumeric ciphers, renowned for their inherent complexity. Breaking such ciphers proved challenging, with success occasionally reliant on fortuitous key conjectures. An approach of searching plain words was applied in this study at first, focusing primarily not on random words but on the vernacular names of plants that with high probability could be in the very text, due to it botanical characteristics of the folios. This emphasis stemmed from the botanical characteristics evident in the folios, due to the complete obscurity of the language employed in the initial stages of deciphering the Voynich manuscript, therefore attention was directed towards Vernacular Latin and Old Greek plant lexicon, which persisted in usage throughout the medieval period in Europe.

This linguistic continuity was facilitated by the extensive transmission of manuscripts dating back to antiquity, notably those attributed to Dioscorides, Pseudo-Dioscorides, Pseudo-Apuleius, Theophrastus, Galen, and Strabo. These texts served as repositories of botanical knowledge, containing a profusion of synonyms within the evolving almost already formed European middle languages of the late Middle Ages, coinciding with the presumed period of composition of the Voynich manuscript. The analysis therefore primarily focused on the identification of specific medieval botanical technical lexicons, which were widely disseminated and utilized during this period.

This method demonstrated a highly positive result in decipherment of the Voynich Manuscript, marking a particularly important moment within this study. For instance, the revelation of the vernacular name for the lavender plant (Lavandula L.) as “lavanda”, proved instrumental in deciphering the alphabet characters, which posed challenges due to the unfamiliar paleography present in the Voynich Manuscript.

For instance, this vernacular botanical name within the Voynich manuscript serves as an essential clue for decrypting the entire text, also offering some insights into the language in which it was written. The initial phase of analysis confirmed the accuracy of this research through detailed examination of the characters ‘T’, ‘A’, ‘V’, ‘N’, ‘D’ within the plaintext of the lavender (Lavandula L.) in the Voynich manuscript.

It became possible to conceptually decipher the decipherment method: deducing the entire alphabet by identifying other plaintext vernacular names of plants containing the characters ‘T’, ‘A’, ‘V’, ‘N’, ‘D’, as well as those vernacular names of plants derived from Vernacular Latin and Old Greek botanical lexicons, which remained in use throughout the medieval period in Europe.

Moreover, the ability to analyse the text of the Voynich Manuscript and recognize which information was relevant, particularly in determining the plaintext vernacular names of plants, made it possible to ascertain at what point there was sufficient information to figure out the decipherment of its cipher. For instance, after determining the botanical lexicon written in plaintext, and thus also establishing the plaintext alphabet characters, the cipher method algorithm used in the Voynich manuscript, along with the ciphered words, was also revealed: it utilized a substitutional encryption algorithm of a polyalphabetic cipher, which was used partially and simultaneously with a mono-alphabetic partial cipher, in addition to completely unencrypted words according to the cipher algorithm's rules in the text. It was also evidently influenced by a transposition cipher found in alchemical texts with double encryption. The Voynich Manuscript exhibits significant differences from ciphers used in the medieval period, due to using partial encryptions of various types.
4.1 Definition of the first plant: lavender (Lavandula L.) vernacular name is the key to a deciphering of the whole Voynich Manuscript.

The botanical specimen referenced in the Voynich Manuscript, lavender (Lavandula stoechas L.), is essential for unlocking the manuscript’s cipher algorithm. Within the scholarly discourse, this proposition suggests that a profound comprehension of the medieval technical vernacular nomenclature associated with lavender (Lavandula L.) serves as a key element for expediting the comprehensive decryption of the manuscript in its entirety.

The key to decipherment and the very first read flower in the Voynich Manuscript is the plant Lavender (Lavandula stoechas L.). Analyzing the text of the Voynich Manuscript on the folia 2r revealed paleographic features of its vernacular name characters: ‘L’ ‘a’ ‘v’ ‘a’ ‘n’ ‘d’ ‘a’ in the plain text. Previous knowledge of this flower, especially its Spanish subspecies, was useful in defining its name in the Voynich Manuscript (fol. 2r). The flower featured in the rebounded Voynich Manuscript, identified as the first plaintext word “Lavanda” on the fol. 2r, corresponding to lavender, served as the key for deciphering the entire manuscript. Paleographic characters of the word “Lavanda”:

| L a v a n d a |
| Lavender |

Fig. 1- Ms.408. fol. 2r Copyright © 2024 Yale University.

(1) The character of the capital letter ‘L’. The majuscule letter ‘L’ resembles two vertical lines with a loop at the top of the second vertical line.

Fig. 2 - Ms.408. fol. 2r Copyright © 2024 Yale University.

(2) The minuscule character of the letter ‘a’ at the first syllable of this word.

The distinguishing feature of the minuscule ‘a’ is its looped form. Instead of a simple closed loop, the top part of the ‘a’ is similar in appearance to the upper part of the numeral 9. There is a tail extending from the descender towards the left.

Fig. 3 - Ms.408. fol. 2r Copyright © 2024 Yale University.

(3) The minuscule character of the letter ‘v’, ‘u’ in the word “Lavanda”.

The letter ‘v’, ‘u’ is a double looped and resembles the numeral 8. It is essential to mention that within the plaintext context, it is important to recognize that during the medieval period, the differentiation between the letters ‘u’ and ‘v’ was not as definitive as it is in contemporary typography. The same character could serve for both the vowel sound /u/ and the consonant sound /w/. The medieval Galician letters u and v were also interchangeable and did not have a significant graphical distinction between them. (Gladyseva, 2019)

Fig. 4 - Ms.408. fol. 2r Copyright © 2024 Yale University.

Some Romance languages have maintained the distinction between a phoneme /v/ and a phoneme /b/. In Romance languages, the letter ‘v’ served to denote both the voiced labiodental fricative sound /v/ and the voiced bilabial approximant sound /w/ in specific contexts. It is imperative to acknowledge that ‘b’ and ‘v’ constitute disparate phonemes in medieval Galician language, mirroring their distinction in contemporary Portuguese. (Gladyseva, 2019)
Cf., after the 1st century AD in reason of depending on Vulgar Latin dialect, consonantal /w/ developed into /β/ , then later to /ν/ in Spain. Therefore, letter ‘ν’ has also the same pronunciation as the letter ‘b’ in Spanish language. (Díez Losada, 2004: 176).

(4) The minuscule character of the letter ‘a’ at the second syllable of this word.

In a single-compartment secretary hand, the minuscule letter ‘a’ is formed within a single enclosed shape. This shape resembles a rounded loop, with a slight curve at the top. The loop shape of the minuscule letter ‘a’ is almost closed, meaning that it has a little extension like the modern cursive form.

(4) The minuscule character of the letter ‘a’ at the second syllable of this word.

In a single-compartment secretary hand, the minuscule letter ‘a’ is formed within a single enclosed shape. This shape resembles a rounded loop, with a slight curve at the top. The loop shape of the minuscule letter ‘a’ is almost closed, meaning that it has a little extension like the modern cursive form.

Fig. 5 - Ms.408. fol. 2r Copyright © 2024 Yale University.

From a paleographic perspective, the variation in the representation of the letter ‘a’ across different syllables can be due to the historical development of scripts. Paleography recognizes also that handwriting styles could vary based on regional preferences and individual scribe habits.

(5) Stylistic ligature of characters ‘n’ and ‘d’ at the almost end of the word “Lavanda”.

It consists of the simple downward stroke shortening [nd]: for characters ‘n’ and ‘d’ at the almost end of the word was quite often used during that period in Romance manuscripts. (A. Gladyseva, 2019)

(6) The minuscule character of the letter ‘a’ at the last syllable of this word.

The distinguishing feature of this minuscule ‘a’ is its looped form. Instead of a simple closed loop, the top part of the ‘a’ is similar in appearance to the upper part of the numeral 9. There is a tail extending from the descender towards the left.

Fig. 7 - Ms.408. fol. 2r Copyright © 2024 Yale University.

3.2 The aspects of the using vernacular name of “Lavanda” for Lavandula stoechas L. in the medieval period.

In general, according to Gerard, Lavandula was the medieval Latin name for lavender, since its later Latin name was lavamdārius (Gerard,1597: 468). But in the very text of the Voynich manuscript, its simpler vernacular name “Lavanda” is used (fol. 2r.). The lexicon "lavanda " as a vernacular name is derived from the Latin gerund of the verb lavare meaning “to wash,” therefore Lavanda means “that which must be washed.” (Gledhill, 2008: 232)

Fig. 8 - Herbario nuovu by C. Durante, 1585. p.307.

In fact, lavender was used by the Romans to perfume their baths (Meunier, 1992), particularly Lavandula stoechas L., and Lavandula stoechas subsp. pedunculata (Mill.), (Lis-Balchin, 2002). Celsus, among other ancient ‘scientists,’ also described its anti-inflammatory properties, since it had the name ‘stoechas’ from the ancient Greek period (στοιχάς) (Celsus, 1938: 532). Cf. also with: Dioscorides III.26, Pliny 20.247, 27.131, Galen Sump. 8.18.39
K.12, 130–1), Marc. Med. 1.106, Isidore 17.9.88. 'Lavanda' entered the lexicon during the 13th century, believed to originate from the Old French word 'lavandre' (Lewis:1879) and as a vernacular name for lavender was used and/or continuing to be used in several languages such as Galician, Spanish, Croatian, Italian, French, Occitan. For example, Mathioli's work (1554) mentions lavandula officinalis, the common Lavender, which is named '[La] Banda' in the medieval Spain.

It is generally known that Lavandula stoechas L., the Spanish lavender or topped lavender (U.S.) also French lavender (U.K.), along with its subspecies, and Lavandula L. (e.g. L. spica and L. latifolia), were considered two separate botanical groups during the ancient period. According to Henry Lyte (1578: 266), who translated Rembert Dodoens' herbal into English in the 16th century, along with William Turner (1548) and John Gerard (1597: 468), the knowledge that these plants are two species of Lavandula was also known. French lavender, for instance, was referred to as "Staechas" or "Sticadove" (Lavandula stoechas L.) in English during the Middle Ages, as its English names in the Middle Ages. As per John Gerard, Stoechas, also referred to as Sticados, received their designation from Dioscorides and Galen, potentially deriving from the Sto切ches, where Stoechas flourished on the islands adjacent to the coast of Marseilles, presently recognized as the Iles de Hyères. (Gerard, 1597: 468),

To ensure the accuracy of the identification of Voynich Manuscript illustration 2r and validate the research findings, it was imperative to locate primary sources from medieval Spanish manuscripts containing references to "Lavanda" or "Lavandula," specifically regarding species of "Stoechas". This was important due to the morphological characteristics of the plant presented on 2r, which differ from those of L. spica and L. latifolia, as well as the research findings regarding the manuscript's origin in Spain and its composition in Medieval Galician, which served as the lingua franca in Spain for an extended period.

Even during the medieval period, when the Voynich Manuscript was likely written, Lavandula stoechas L. had numerous vernacular synonyms. For example, in Spanish, it was known as “Cantueso” and “Rojmarinho” (Gerard, 1597: 468). Manuscripts at Salamanca University also reveal other vernacular names such as “Sticados” (MSS 2262, 1501: fol 108r), while those at the National Library of Spain mention nomenclature like “Sesticados” and “xalca” in medieval Galician (BNE MSS/3338, 1401-1500: 136-147v).

It must be admitted that the use of the medieval nomenclature "Lavanda" or "Lavandula," specifically regarding species of “Stoechas”, was not frequent in Spain but also in other countries during that period. However, it was not exceptional, as evidenced by its appearance in the manuscript "Libro del arte de las comadres" (1541) written by Damián Carbón: “dracma media laundule sticados/ polegij/ montani/ camomilloti” f. 38r; “Recipe laundule sticados,” f. 58r; “Despues hagan le batos y sufumigos con cosas apropiadas con artemisal/ con assensio/ laundula” f. 103v (Carbón: 1541).

Since Carbón’s manuscript was written in Old Spanish, it also extensively used other plants' vernacular nomenclatures of Latin origin, known from medieval reworkings of herbal manuscripts, such as those of Dioscorides, Pseudo-Dioscorides, Pseudo-Apuleius, and so on. This suggests the use of such nomenclatures of Latin origin as part of the medieval "scientific" lexicon in the pre-Linnaean period.

Moreover, it shows that laundule sticados (or 'lavandule sticados") was an eligible existing name in the botanical-medicinal lexicon of medieval Spain.

3.3 The mention of particular species within the Lavandula L. genus on folio 2r of the Voynich Manuscript.

The specific subspecies of Lavandula stoechas L. referenced in the Voynich manuscript is identifiable through the vernacular name of the plant in the deciphered text and the depiction of red stems supporting Sticadove pale pink flowers with its fragrant petals. This closely resembles the characteristics of Lavandula stoechas L., particularly Lavandula stoechas subsp. pedunculata (Mill.) and Lavandula pedunculata subsp. latifolia. The contemporary vernacular Spanish name for Lavandula stoechas subsp. pedunculata (Mill.) is el espliego (spain.inaturalist.org, 2024), borrowed etymologically from Aragonese espligo, from Latin spicaulum, meaning 'little spike' (Coromines & Pascual 1984: 750).

Due to specifications of the coding algorithm, which are based on the logical characteristics of a substitutional encryption algorithm of a polyalphabetic cipher (rather than on mathematics), since it was used partially, and also simultaneously with a mono-alphabetic partial cipher, this resulted in completely unencrypted plain words according to the cipher algorithm's rules in the text.

<table>
<thead>
<tr>
<th>Lavender</th>
<th>n def.art.</th>
<th>Spikes</th>
</tr>
</thead>
<tbody>
<tr>
<td>L a v a n d a</td>
<td>a a P i c o s</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 9 - Ms.408, fol. 2r Copyright © 2024 Yale University.

(1) "Lavanda" n. identified as a feminine singular noun, medieval vernacular name for "lavender."

(2) "a" is identified as the determinate feminine singular article. An important key to determining the language in which the Voynich manuscript was written was the presence of Medieval Galician articles, including the specific singular masculine definite article ‘o’ and singular feminine article ‘a’. For centuries, scholars thought that an ‘o’ symbol was a zero or a coded symbol, but in reality, it represents the singular masculine definite article ‘o’, which
was not encrypted at all. According to this research, the Voynich manuscript was written in the Medieval Galician language (Gladyseva, 2018; 2019; 2020).

It is essential to examine some medieval aspects concerning the unstable use of definite articles in this manuscript. Prior to the noun "Lavanda," its definite feminine singular article "a" is omitted. Instead, it follows the noun, mentioned alongside the word "Picos" clearly indicating a definite feminine singular association with the noun's feminine suffix.

(3) "Picos" n. masc. plural. It derives from medieval Galician pico, from Vulgar Latin *piccus, m.; 1. sharp point; 2. spike (Latin spīca), since in Galician it also means thorn as espiña, and spike, synonym to espliego. According to medieval Galician lexicón “Vocabulario clasificado de los folios gallegos de la Historia Troyana”, the word “Picos” means “pico – parte puntiaguda; and translates in English as a sharp point” (K. M. Parker, 1977). In contemporary period, this lexicón remains employed for botanical purposes within the Galician linguistic domain, cf. Pico: hoja del pino, in Eng.: a leaf of pine. (Vázquez Lopez al., 2000: 727)

Due to the detailed descriptions of the plant in the text and the illustrations in the Voynich manuscript, it should not be confused with other species (e.g. Lavandula angustifolia (synonymous: Lavandula spica L.) and Lavandula latifolia (synonymous: Lavandula spica subsp. latifolia)), which were also well translated into English by Lyte (1578: 265) from Dodoens herbal and described by Gerard using medieval English lexicón, such as ‘Lavander’, ‘Lavander Spike’ (Gerard, 1597: 468).

In medieval scholarly discourse, the Middle English name ‘Sticadove’ was also accompanied by the designation ‘a spike,’ as Gerard differentiates ‘Jagged Sticados’ with spiky characteristics: “on the top of the stalks do growe spike flowers” (Gerard, 1597: 468). Therefore, it is of high probability that in this case ‘picos’ also delineates this characteristic, since the Spanish translation of spike is ‘espliego’, which is synonymous with the medieval Galician ‘pico’. The etymology of ‘espliego’ traces back to the Latin word spiculum, meaning ‘spike’ (Corominas & Pascual 1984: 750). It means that due to the commonalities of the flowers, that are gathered in dense terminal spikes of quadrangular section in plant morphology between the Stoechas species and ‘Spike’ (Lavandula spica L.) species, they also attributed not all, but the same vernacular names. Moreover, the vernacular names for ‘Sticadove’ and ‘Spike’ are frequently the same, as in Spanish; for example, both plants are known as ‘espliego’, ‘espligo’, ‘espliego’, or ‘Lavanda’. In Galician, they are both known as ‘espliego’, ‘lavanda’ (Anthos, 2012).

These results closely resemble the possibility, since the Voynich manuscript contains an illustration and description of a subspecies of L. stoechas L. that was known in medieval times only in general as Stoechas. Given the fertility of Spanish soil, it could be, for example, Lavandula stoechas subsp. pedunculata (Mill.); Lavandula pedunculata subsp. lusitanica; Lavandula stoechas L. subsp. stoechas; or Lavandula stoechas subsp. lusieri (Rozeira); since there was no understanding of such nuances in the medieval period as these subspecies of the species Lavandula stoechas L. However, according to Rozeira (1949), the Lavandula stoechas subsp. pedunculata (Mill.), without its subspecies, has long been recognized by Dodonaeus (1568), and also Lobelius (1570) (Gagnus, 1826). Despite this, in Species Plantarum (1753), Linnaeus did not assign a specific epithet to Lavandula stoechas subsp. pedunculata (Mill.); he simply recognized it as var. s under Lavandula stoechas L., therefore it had no nomenclatural status. The distribution of subsp. pedunculata is found in Spain and northeastern Portugal, while subsp. lusitanica and subsp. sampaiana are found in southwestern Spain and Portugal (Nieto Ojeda, 2006). These subspecies are highly polymorphic. For example, Lavandula stoechas subsp. pedunculata (Mill.) has a long peduncle of the inflorescence longer than twice its length, and its botanical name. The long peduncles that support the spikes are devoid of leaves, similar to the illustration in the Voynich Manuscript f. 2r. This subspecies forms a highly branched bush that reaches 80 cm in height, with narrow, lanceolate lanose leaves that often form a fascicle. The entire plant gives off a pleasant and penetrating camphorous smell. (Nieto Ojeda, 2006).

These results suggest that the Voynich manuscript f. 2r contains a description of Lavandula stoechas L., specifically Lavandula stoechas subsp. pedunculata (Mill.), with a possible variation to Lavandula pedunculata subsp. lusitanica due to the pale colour of the flowers and the red stem. Notwithstanding, f. 2r pertains to the Lavandula stoechas L. species.

3.4 The potential inclusion of an additional species from the Lavandula L. genus depicted on folio 2r of the Voynich Manuscript.

According to scholarly observations, in which no decipherment of the Voynich Manuscript has been made, but rather suggestions, the texts are segmented into paragraphs, ostensibly indicating a shift in topic (Bowern & Lindemann, 2021: 16). However, this innovative decipherment research presents a different view. It proposes that the paragraph divisions on botanical folios rarely reveal different dimensions of the same subject, highlighting rare nuances. For example, only a few folios detail several similar plants in consecutive paragraphs, each discussing distinct characteristics of plants within the same genus, it means, within the same topic. The second paragraph begins with the words:

<table>
<thead>
<tr>
<th>Lavender</th>
<th>grand</th>
</tr>
</thead>
<tbody>
<tr>
<td>L a v a n d [a]</td>
<td>g r a n d</td>
</tr>
</tbody>
</table>

Figure 10: Ms.408, fol. 2r Copyright © 2024 Yale University.
(1) Lavanda – n. f. sing., lavender
(2) grand – adj. grand

According to K. M. Parker (1958): Vocabulario de la Crónica Troyana. grand (also grande and grand) means adj. grande; large in medieval Galician. (Parker, 1958)

In consideration of the presence of multiple plant descriptions within the same genus on other folios of the Voynich manuscript, the absence of syntactic punctuation in the text introduces ambiguity, allowing for another interpretation. Following a comprehensive analysis of the entirety of the second paragraph on folio 2r, it becomes conceivable to entertain the notion that the text may either extend the description of the plant introduced in the first paragraph or introduce an additional species within the same genus.

If the second paragraph of folio 2r introduces an additional species within the same genus, it likely describes the broadleaved lavender, also known as Portuguese lavender. This flowering plant is native to central Portugal and Spain. In Spanish, it is referred to as "gran lavanda," as noted in other languages such as Italian "$\text{lilium lavanda}" and French "$\text{Grand Lavande}". (Lamendin, 2007. (projectnoah.org)

Its contemporary scientific name is Lavandula latifolia, with the synonym Lavandula spica subsp. latifolia. Additionally, in English, it is also known as Spike lavender. These findings suggest that the Voynich manuscript may also contain descriptions of this plant.

Nevertheless, given the absence of syntactic punctuation, there remains a possibility that the text in the second paragraph of folio 2r extends the description initiated in the first paragraph. This is attributable to the morphological largeness of Lavandula stoechas L. and its subspecies. For instance, Lavandula stoechas subsp. pedunculata (Mill.) is also identified by vernacular names such as Lavanda Largo. (spain.inaturalist.org)

3.4 Validating the accuracy of identifying the letter 'L' in additional plain texts, including vernacular plant names in the Voynich manuscript.

Because of the coding algorithm's specifications, which rely on the logical properties of a substitutional encryption method known as a polyalphabetic cipher, rather than on mathematical principles, and due to its partial utilization alongside a mono-alphabetic cipher, this led to entirely unencrypted plaintext words in accordance with the cipher's rules within the text.

It provided the opportunity to search for other plaintext words, especially concerning vernacular plant nomenclature, which could also be derived from medieval Latin, given the linguistic uncertainty surrounding the language used in the Voynich manuscript during the initial stages of this research.

I. Validation example:

In the second paragraph of folio 2v of the Voynich manuscript, it was possible to find plaintext of the vernacular name of the plant Nymphaea alba L., commonly known as the white waterlily, that was accompanied with the plaintext 'grand,' serving as a medieval Galician adjective. grã.

<table>
<thead>
<tr>
<th>Waterlily</th>
<th>grand</th>
</tr>
</thead>
<tbody>
<tr>
<td>L i r o r</td>
<td>grã.</td>
</tr>
</tbody>
</table>

Fig. 11 - Ms.408. fol. 2v Copyright © 2024 Yale University.

(1) 'Liror' – n. m., meaning waterlily, which also exhibits the characteristic medieval Galician singular suffix -or. It serves as a synonym for 'lilio,' as seen in sources such as the Cantigas de Santa Maria of the 13th century (see W. Mettmann, 1972: 24.57 'liro'; 24.2 'lilio'). The term originates from Latin illium and Ancient Greek 'λιένον,' and is used in contemporary Spanish and Galician as Lirio de agua. Cf.: 'lilio o de nenofari' in the medieval Tratado de cirugía (Cauliaco, Nacional INC/196, 1498: f. 131v). This plant is also mentioned in the medieval Spanish-Galician manuscript MS. 3338, where it is listed as 'lírio de agua,' and is also listed under the chapter titled 'Flor de las yervas,' containing vernacular plant names alongside synonyms in medieval Galician, such as 'lazon,.i. yerva del lago,.i. unguella, cubulama,.i. nenofari' (MS.3338: 143va). It is no coincidence that the marginal note "lir" above the word "Liror" suggests that someone associated the origin of this plant's name with the Old French vernacular term "lis," which was derived from the earlier forms "lil" in the nominative singular and accusative plural, ultimately tracing back to Latin lillium. However, it is more likely that the individual who made the marginal note intended to reference the medieval Occitanvernacular name for this plant: "Liri" or "Leri."

(2) grã - adj. (adv.) grand. 1. Applied to size, "GRÄ", del lat. grandis, according to O léxico dos Mirages de Santiago of M. C. Barreiro (1985).

The plain text analysis revealed an interesting observation regarding the medieval Galician vernacular name for a particular plant in the first paragraph. This finding suggests that the master of the Voynich manuscript employed synonyms to refer to the same plant, likely to avoid repetitive usage of the same name.

<table>
<thead>
<tr>
<th>Waterlily</th>
<th>*in medieval Galician</th>
</tr>
</thead>
<tbody>
<tr>
<td>L a [z] o n</td>
<td>(d)</td>
</tr>
</tbody>
</table>
Fig. 12 - Ms.408. fol. 2v Copyright © 2024 Yale University.
(1) L a [z]o n (d) – n. s. medieval Galician vernacular plant name according to 'Flor de las yervas,' containing vernacular plant names alongside synonyms in medieval Galician, such as 'Lazor,.i. yerva del lago,.i. unguesella, cubalina, .i. nenufar' (MS.3338: 143va). In medieval manuscripts, the omission of various letters within words is not uncommon, with the letter "z" being just one example and does not constitute an anomaly. Furthermore, the addition of the specific suffix "d" to the end of a word is a characteristic feature of medieval Galician language.

2. Validation example:

However, the master of the Voynich manuscript also integrated the word "liro" as a synonym for the term "flor," "flower," which was also utilized in medieval Galician. This likely served the purpose of describing it as part of the plant's morphology or to obscure the text for unfamiliar readers. Moreover, the master avoided repeating essential words such as vernacular plant names or the term "plant" throughout the entirety of the manuscript's folios. For example, on folio 4v, the presence of the word "liroa" as a botanical morphological lexicon, instead of "flor," is notable in the description of Gentiana pneumonanthe, L.'s flower in the text used by the master of the Voynich Manuscript.

[flower: in this botanical context, it serves as a synonym for 'flor']

Fig. 13 - Ms.408. fol. 4v Copyright © 2024 Yale University.
(1) 'Liroa' – n. f. s., (see W. Mettmann, 1972: 24.57 'liro'; 24.2 'lilio'). Its meaning is "flower" in this morphological context, serving as a synonym for 'flor' also throughout the Voynich Manuscript. Moreover, notably, it exhibits the most characteristic feminine suffix –oa for medieval Galician language.

3. Validation example:

As a validation of decipherment accuracy, observation on fol.32r reveals consistent palaeography for the letter 'L' in the plaintext, as seen in previous instances like the name 'Lavanda.' (fol. 2r). The description of Hesperocodon hederaceus L. (synonym Wahlenbergia hederacea L.), also known as the ivy-leaved bellflower on fol. 32 r of the Voynich Manuscript, begins with the medieval term 'Lirios.', which can be translated simply as 'lily' or possibly with another etymological derivation from Old French 'Liere.' It's noteworthy that numerous Spanish and Galician plants are assigned contemporary vernacular names like 'Lirio.' For instance, nearly 440 different plants, including those with morphologically some kind similar bluebell flowers such as Romulea bulbocodium L., Epilobium hirsutum L., and Paradisea liliastrum L. (Anths, 2024)

['lily' or possibly with another etymological derivation from Old French 'Liere']

Fig. 14 - Ms.408. fol. 32r Copyright © 2024 Yale University.
(1 a) lirio s. m., from medieval Galician liro, lilio, from Latin illium, from Ancient Greek ἵςίπυον (leirion), c.f. lirio: 24.2 from Cantigas de Santa María de Afonso X, o Sábio. (Mettmann, 1972)

(1 b) or possibly with another etymological derivation from Old French 'Liere.' Due to its leaves' similarity to the Ivy plant, the master of the Manuscript also mentions it in the first paragraph, where the fourth plain word variant is 'Liera.' (fol. 32 r).

[Ivy: cf. Old French: Liere or medieval Galician laera.]
Liera derives from Old French ‘lire,’ formed by the elided article ‘l’ and ‘iere’ (m), itself from Latin ‘hedera’ (f). This corresponds with medieval Galician: ‘éra,’ ‘hera,’ meaning ‘hiedra’ cf. Livro Enrique: ‘toma a flor do apoio e do salgueiro e dous grãos da era’ (Lorenzo, 1977), with the possible addition of the definite f. article ‘l’a’ (Barreiro, 1985), as ‘l+era’ in a similar manner to Old French.

The late specific epithet, hederacea, –ea, in Hesperocodon hederacea L. (synonym Wahlenbergia hederacea L. derives from classical Latin and is a declension of the adjective hédéracêus, which means “like ivy; similar to ivy.” (Online Latin Dictionary, 2023)

4. Validation example:

The next validation example is quite evident and observable on f30r in the vernacular plant name of Convallaria majalis L., also known as Lily of the Valley, which remains referred to in contemporary Galician botany as ‘lirio do val’.

The master of the Voynich manuscript initially used the same plain name in a medieval Galician manner, and also likely adding the feminine suffix ‘–oa.’ at first, resulting in ‘lirioa’. However, this was later changed to a more specific medieval Galician feminine suffix, ‘–za,’ resulting in ‘liriza’, since he used feminine article ‘a’ at the beginning.

Fig. 16 - Ms.408. fol. 30r Copyright © 2024 Yale University.

(1) a f. art. det. The article is a noun presenter, indicating also its gender. (Barreiro, 1985)


Given that the entire text of the Voynich manuscript has already been deciphered, based on the detailed specifications provided in the plant description, it is unequivocally clear to affirm that this plant corresponds to Convallaria majalis L.

5. Validation example:

The next plant was well-known in the medieval period and had a Latin vernacular name, Lunaria, derived from ‘Lūna,’ due to the white-silver color of the fruit similar to that of the Moon. Various types of Lunaria are described in different manuscripts from the medieval period, as it was utilized in alchemy. For example, a manuscript from the 15th century at the University of Pennsylvania LJS 419, titled “Erbario,” contains entries on fols. 78v, 79r, 79v, and 80v specifically about different subspecies of the Lunaria plant. The variation between types is evident in the shape of the leaves, which can resemble for e.g. a half moon, as well as the seedpods resembling a full moon. The 15th-century manuscript MSS/3338, ‘Suma de la flor de cirugía,’ also makes reference to this plant in Spanish as ‘Luna’; ‘olio de luna.’ (MSS/3338, XV c. f. 187r)

The mention of Lunaria annua L. on folia 13v of the Voynich manuscript serves as essential evidence for the accuracy of defining medieval Galician as the language of the Voynich Manuscript. Besides proof specific medieval Galician articles and suffixes, it also provides proof of the medieval Galician lexicon, as the Latin vernacular name Lunaria, meaning “moon-like” in reference to the decorative seedpods, was directly translated into medieval Galician by the master of the Voynich Manuscript as ‘Lūair’.

Fig. 17 - Ms.408. fol. 13v Copyright © 2024 Yale University.

(1) ùa n. f. moon, according to Vocabulario clasificado de los folios gallegos de la Historia Troyana. C.f. a ùa, 281.12, 278.31, 343.18. (Parker, 1977)

The name Lūair mentioned here includes the essential medieval Galician suffix “-ir.” Cf.also ùa n. moonlight: II 214.8, I 300.8. (Parker, 1958)

The vernacular plant name "Lunaria" and its synonyms, such as "Plata," are still used in contemporary botany for Lunaria annua L. For instance, in Galician, it’s referred to as "Lunaria," in Spanish as "Plata" or "Lunaria," and in Portuguese as “Lunária.” The fourth plain word in the first paragraph also represents the morphological term “lios” for flowers, instead of “flor.”
(1) 'Liros’ – n. f. pl., (see W. Mettmann, 1972: 24.57 'liro'; 24.2 'lilio'). The meaning is 'flowers' in this morphological context, serving as a synonym for 'flor' also throughout the Voynich Manuscript.

4. The real plaintext alphabet of the Voynich Manuscript.

The problem with previous research was the determination of the plain text alphabet characters in the Voynich Manuscript, which seemed to exhibit no resemblance to any known writing system. It is obvious, even trying to read each line of the Voynich manuscript with an IBM punch card machine by the "First Study Group," led by cryptographer William F. Friedman in the 1940s, was unsuccessful. Moreover, the European Voynich Alphabet (EVA) is in substance incorrect, leading to failure in any attempts to decode based on this alphabet:

Edith Sherwood’s interpretation of the botanical texts holds particular interest due to her recognition that they were composed in a European language or its dialect. Alongside Ethel Voynich (1864–1960) and Ellie Velinska, Sherwood Ph.D. made a significant contribution to the field by accurately identifying certain plants visually, despite discrepancies in the accompanying alphabet:

The only reason is the incorrect decoding of characters in the Voynich manuscript and a misunderstanding of the language in which it was written. (A. Gladyseva, 2019)

Based on the innovative research results presented in this article, involving only the determination of plaintext letter characters and the provision of these five validation examples, it is possible to state the true plaintext alphabet of the Voynich Manuscript:

5. Conclusion

The present study emphasizes the importance of this research as the first deciphered data of the Voynich Manuscript. The encryption code of the Voynich manuscript that made decryption more complicated over the centuries was the use of specific cipher algorithm of polyalphabetic encryption that was used partially and simultaneously with monoalphabetic encryption (Gladyseva, 2020). The implementation of this specific cipher algorithm resulted in the transformation of numerous words into plaintext, implying that they were effectively left unencrypted. The plaintext, as it is seen, is very clear and
obvious, requiring no additional proof of decipherment, due to the coding system of the Voynich Manuscript is partial, where plaintext is seamlessly integrated into the text without alteration. Moreover, the plaintext facilitates the determination of the language in which it was written, specifically medieval Galician, evident not only through its lexicon but also through the usage of medieval Galician articles and specific suffixes.

The results established the correct alphabet characters of the Voynich Manuscript, since suggestions like EVA (European Voynich Alphabet) and other alphabets, which were not based on the primary text of the Voynich Manuscript, continuously led to erroneous attempts at decipherment for almost a century, resulting in completely inaccurate scientific conclusions, whether made by humans or IBM. Therefore, the findings of this research represent the only deciphered data, first presented to the audience.

The simple analysis of phytomorphs' similarities with herbarium specimens conducted by other researchers, generally resulted in erroneous categorization of plant species. This is largely attributed to the challenges in determining the species, given the mostly partial or exaggerated depictions, often of immature plants, in the Voynich Manuscript.

Therefore, the methodology of decipherment in this research was appropriately chosen, as it is the only study of the Voynich Manuscript primarily and generally based on its linguistic source—the lexicology. For example, the name "lavanda" was primarily deciphered and read directly within the text of the Voynich Manuscript.

The data obtained thus allows to suggest the high competence of the master of the Voynich Manuscript in languages and medieval nomenclature of the plants and also tried in some cases to translate some vernacular names from Old Greek, Latin, Old Spanish, Old French, Occitan origin into medieval Galician with using specific medieval Galician suffixes, and adding articles in a proper manner. However, it important to admit of existing of a lot medieval Galician vernacular names of plants and also tried in some cases to translate some vernacular names from Old Greek, Latin, Old Spanish, Occitan origin into medieval Galician with using specific medieval Galician suffixes, and adding articles in a proper manner. However, it important to admit of existing of a lot medieval Galician vernacular names of plants and their synonyms due to natural variety of plants in Spain area and a lot of medieval primary sources data lexicon on them, for e.g. ‘Flor de las yervas,’ containing vernacular plant names alongside synonyms in medieval Galician and Old Spanish. (MS.3338: 143va). The manner of description of plants reminds those described by Dioscorides, Apuleius, but mostly to Theophrastus’ works, due to the emphasis made by the master of the Voynich Manuscript on the rhizotomos system, with detailed descriptions of the roots. Additionally, it suggests more references to the lost medieval period manuscript of Crateus, known for its accent on root systems, according to his work “De Materia Medica” a herbal encyclopedia of the 1st century BCE.

Acknowledgements

Dedicated to My Beloved Mother, Galina Gladyshevka, for her love, support, and for constantly believing in me.

I would also like to acknowledge Vilnius University and the academic members of the Department of Romance Studies and the Faculty of Philology in general: Doc. Dr. Miguel Villanueva Svensson, Prof. Dr. Jurgis Pakerys, Prof. Dr. Aksel Holvoet, Asist. Dr. Maria Sebastia Saez, and Doc. Dr. Aleksej Burov.

I’d also like to extend my gratitude to my family for all the support: G.D. M. Gladyshev, J.Gladyshev, E.Gladysheva, M.Avery and A.Luck.

References

Beinecke MS 408, Cipher Manuscript (Voynich Manuscript), General Collection, Beinecke Rare Book and Manuscript Library, Yale University.
Cipher Manuscript (Voynich Manuscript). Beinecke MS 408. General Collection, Beinecke Rare Book and Manuscript Library, Yale University.


Dodoens, R. *Florum, et coronariarum odoratarumque nonnullarum herbarum historia* (p. 179).

Dodoens, R., & Lyte, H. (1578). *A nieue herball, or, Historie of plantes wherein is containyd the vhole discourse and perfect description of all sorts of herbes and plantes, their diuers and sundry kindes ... and that not onely of those whiche are here growynge in this our coutrie of Englane but of all others also of forrayne realmes commonly used in physicke*. Gerard Dewes.


Gilbertus, A. (1510). *Compendium medicine Gilberti anglici tam morborum universalium quam particularium nondum medicis sed et cururgicis utilissimum*.


Isidore of Seville. (1599). *Etymologiae* (Book 17, Chapter 9, Section 88, J. de Grial, Ed.). Madrid.


Manuscript BNE MSS/3338 Sesticados and xabca, National Library of Spain, fols. 136r-147v), 1401-1500

Manuscript, LJS 419, "Erbario," Lunaria plant, fols. 78v, 79r, 79v, 80v, the University of Pennsylvania, the 15th century.

Manuscript MSS 2262, Sticados, Salamanca University, fol 108r, 1501

Matthioli, P. A. (1554). Medici Senensis Commentarii, in Libros sex Pedacii Dioscoridis Anazarbei, de Materia Medica, Adjectis quâm plurîmis plantarum & animalium imaginibus, eodem authore. Also known as Commentarii.


Oxford English Dictionary. (1989). Note however that Upson and Andrews refer to research on bathing in the Roman Empire, and state that there is no mention of the use of lavender in works on this subject (2nd ed.). Oxford.


