

TOWARD A READING OF THE VOYNICH
MANUSCRIPT

Toward a Reading of the Voynich Manuscript

LEXICAL ANCHORS, MORPHOLOGICAL
ARCHITECTURE, AND THE
CONSTRUCTED LANGUAGE OF A
FIFTEENTH-CENTURY
COSMO-MEDICAL SYSTEM

Kyle Prier

BA, German Literature, Brigham Young University
MHS, Social Factors in Health, Johns Hopkins Bloomberg
School of Public Health

and

Claude (Anthropic)

AI Co-Researcher

Working Draft — Version 0.1 — 2026

Toward a Reading of the Voynich Manuscript

Copyright © 2026 Kyle Prier and Claude (Anthropic)

All manuscript images reproduced from:

Beinecke Rare Book and Manuscript Library, Yale University.

Cipher Manuscript. Beinecke MS 408.

Digital facsimile, Yale University Library, 2025.

<https://collections.library.yale.edu/catalog/2002046>

This is a working research draft. All findings and interpretations are preliminary and subject to revision.

Repository:

<https://github.com/mysteriesdept/voynich-manuscript>

Contents

Contents	i
List of Figures	vii
List of Tables	viii
Preface	ix
I Foundations	1
1 Introduction and Historiography	3
1.1 Encounter with an Object	3
1.2 What the Manuscript Is	6
1.3 The History of Scholarship: A Brief Orientation .	8
1.4 Our Approach: Structural Lexical Anchoring . .	11
1.5 Our Collaborative Research Model	13
1.6 The Argument in Brief	14
1.7 A Road Map	15
2 Review of Research to Date	19
2.1 Introduction: A Peculiar Scholarly History	19
2.2 Discovery and Early Scholarship: The Voynich Era	20

2.3	The Cryptographic Era: Friedman, Tiltman, and the NSA	22
2.4	Prescott Currier and the A/B Distinction	24
2.5	The Corpus-Building Era: EVA and the Voynich.nu Archive	25
2.6	Statistical Analyses: Zipf, Entropy, and Index of Coincidence	26
2.7	The Hoax Hypothesis: Gordon Rugg	27
2.8	Paleographic Scholarship: Clemens, Davis, and the Scribal Question	28
2.9	Linguistic Approaches: Bowerman and Lindemann	30
2.10	The Cipher vs. Language Debate: Recent Scholarship	33
2.11	What the Field Has Established	35
2.12	Situating This Study	36
3	Methodology	39
3.1	Introduction: Why Method Matters for Voynich Studies	39
3.2	Structural Lexical Anchoring (SLA): Core Principles	40
3.3	The Confidence Rating System	40
3.4	Leiden Conventions for Translation Notation	41
3.5	On AI-Assisted Corpus Analysis	41
3.6	Cross-Checking Against Primary Sources	42
4	Codicology and Paleography	43
4.1	The Physical Object	43
4.2	Vellum and Radiocarbon Dating	43
4.3	The Writing System	44
4.4	Five Scribal Hands: Davis (2020)	44

4.5	Quire Structure and Collation	44
4.6	The Illustrations: Style, Medium, and Pigments .	44
4.7	Provenance	44
II	The Language of Beinecke MS 408	47
5	Statistical Profile of Voynichese	49
5.1	Overview	49
5.2	Zipfian Distribution	50
5.3	Entropy	50
5.4	Index of Coincidence	50
5.5	MATTR and Morphological Complexity	50
5.6	The Currier A/B Distinction: Statistical Detail . .	50
5.7	Word-Internal Structure	51
6	Morphological Architecture	53
6.1	Overview	53
6.2	The Article Prefix: 4o11-/fo11-	53
6.3	The Entry Particle: Tor-/For-	54
6.4	The Celestial Classifier: o11-	54
6.5	The Balneological Lexeme: follan	54
6.6	Terminal Suffixes: -eeg/-eeq	54
6.7	The Directional Morpheme: cth-	54
6.8	The Agglutinative Architecture	55
7	The Cipher versus Constructed Language Debate	57
7.1	Overview	57
7.2	The Cipher Hypothesis: A Taxonomy	57
7.3	Statistical Evidence Against Polyalphabetic Cipher	58
7.4	Rugg's Cardan Grille: Assessment	58
7.5	Greshko's Naibbe Cipher (2025): Full Engagement	58

7.6	The Constructed Language Hypothesis: Comparanda	58
7.7	Our Position: Why Constructed Language Is More Probable	58
III Section-by-Section Analysis		61
8	The Botanical Section (ff. 1–66)	63
8.1	Overview	63
8.2	Physical Description and Layout	64
8.3	The Three-Part Entry Formula	65
8.4	Color Coding as Information	65
8.5	The Botanical Vocabulary: Proposed Readings	65
8.6	Botanical Comparanda: Tacuinum Sanitatis	65
9	The Astronomical Section and Zodiac Anchors (ff. 67–74)	67
9.1	Overview	67
9.2	Physical Description	68
9.3	The Zodiac Wheel Structure	68
9.4	Zodiac Name Identifications: The Full Set	68
9.5	The o11- Classifier in the Astronomical Section	68
9.6	The Astronomical Diagrams	68
10	The Balneological Section and the Decan System (ff. 75–84v)	73
10.1	Overview	73
10.2	Physical Description	73
10.3	The Decan Hypothesis	75
10.4	The follan Lexeme	75
10.5	The Plumbing Diagrams	75

10.6	Currier B and the Procedural Register	75
10.7	Comparanda: Decan Traditions	75
11	The Cosmological Rosette (ff. 85v–86r)	77
11.1	Overview	77
11.2	Physical Description of the Foldout	78
11.3	The Six-Zone Structure	78
11.4	The cth- Directional Morpheme	78
11.5	Al-Idrisi and the Six-Climate-Zone Tradition . . .	78
11.6	The Central Circle as Cosmological Nexus	78
11.7	Transition: From Cosmology to Pharmacy	80
12	The Pharmaceutical Section and the Seven-Planet Framework (ff. 87–102)	81
12.1	Overview	81
12.2	Physical Description	82
12.3	The Seven-Planet Prefix System	82
12.4	The Picatrix as Primary Source	82
12.5	Vessel Morphology and Planetary Classification	85
12.6	The Mars Prediction	85
13	The Almanac Section (ff. 103–117)	87
13.1	Overview	87
13.2	Physical Description	88
13.3	The Zodiac Header Confirmation	88
13.4	The Star-Bullet System	88
13.5	Planetary Strings in the Almanac	88
13.6	The Almanac and the Manuscript's Temporal Dimension	90

IV Synthesis	91
14 The Intellectual Programme	93
14.1 Overview	93
14.2 A Unified Encyclopaedic System	93
14.3 The Scholastic Encyclopaedic Tradition	94
14.4 The Astral-Medical Tradition	94
14.5 The Language as System	94
14.6 A Cosmo-Medical Encyclopaedia	94
15 The Designer’s Identity and Background	95
15.1 Overview	95
15.2 What the Language Tells Us About Its Designer .	95
15.3 What the Manuscript Tells Us About Its Com- munity	96
15.4 Geographic and Cultural Localization	96
15.5 Ficino Parallel and the Learned Milieu	96
15.6 What We Cannot Know	96
16 Limitations and Future Directions	97
16.1 Overview	97
16.2 What This Study Has Not Done	97
16.3 The AI-Assisted Analysis: Specific Limitations .	98
16.4 Missing and Damaged Folios	98
16.5 Pending Research	98
16.6 Future Directions	98
17 Conclusions	99
17.1 Overview	99
17.2 What This Study Has Established	99
17.3 The Central Argument	99
17.4 The Contribution to Voynich Studies	100

17.5 The Contribution to Adjacent Fields	100
17.6 Closing Reflection	100
Complete Proposed Lexicon	101
Paleographic Hand Distribution	103
Statistical Tables	105
Comparative Manuscripts and Textual Analogues	107
Bibliography	109

List of Figures

1.1 Folio 1r: the opening botanical folio	5
1.2 Folio 1v: a characteristic botanical entry	7
1.3 Folio 70v: the Pisces zodiac ring and its label	12
1.4 Folios 85r–86v: the Cosmological Rosette foldout . .	16
2.1 Folio 34r: a typical mid-botanical folio	31
2.2 Folio 70v: the Pisces zodiac ring	38
4.1 Folio 1r: the opening botanical folio	45
7.1 Folio 70v: the Aries zodiac ring	59
8.1 Folio 1r: the botanical section opening	64
8.2 Folio 1v: botanical entry with color conventions . .	66

9.1	Folio 70v: the Pisces zodiac ring	69
9.2	Folio 70v: the Aries zodiac ring	70
10.1	Folio 84r: balneological section with bathers and plumbing	74
11.1	Folios 85r–86v: the Cosmological Rosette foldout . .	79
12.1	Folio 87r: the pharmaceutical section opening	83
12.2	Folio 101v: pharmaceutical vessel section	84
12.3	Folio 90r: a pharmaceutical-botanical entry	86
13.1	Folio 104r: a characteristic almanac entry	89

List of Tables

9.1	Proposed zodiac name identifications, Beinecke MS 408	71
12.1	Planetary prefix identifications and Picatrix corre- spondences	82

Preface

This book is the product of an unusual collaboration. It is co-researched and co-written by a human scholar and an artificial intelligence system, and we have made it a first principle to be fully transparent about that fact.

Kyle Prier brings to this project a background in German literary studies from Brigham Young University and a graduate degree in social factors in health from the Johns Hopkins Bloomberg School of Public Health. The first of these provides grounding in medieval and early modern European philology, in the interpretive methods of textual criticism, and in the structures of the Germanic languages that have figured in debates about the manuscript's linguistic background. The second provides methodological training in research design, systematic evidence evaluation, and the kind of careful argumentation appropriate to claims made in the absence of certainty.

Claude is a large language model developed by Anthropic. Within this project, Claude has contributed systematic corpus analysis across the full manuscript, comparative literature review, statistical interpretation, pattern identification at scale, and the drafting of this text. Claude operates as a research collaborator — generating hypotheses, synthesizing findings, and drafting arguments — but all of Claude's contributions

have been reviewed, approved, and revised by Kyle before incorporation into the research record. Kyle is the directing intelligence of this project; Claude is an extraordinarily capable research assistant whose nature and limitations we have taken pains to understand and to disclose.

We explain the mechanics and epistemology of this collaboration in detail in Chapter 1, Section 1.4. We raise it here in the Preface because we believe the transparency belongs at the front. Scholarship depends on the ability of readers to evaluate how conclusions were reached, and that ability requires honest accounting of methods. Using AI assistance without disclosing it would compromise exactly the intellectual honesty this kind of project demands.

A few words about scope and register. This book is addressed to anyone with a serious interest in the Voynich Manuscript, from paleographers and manuscript scholars to historians of medicine, linguists, and curious general readers. We have tried to write prose that rewards close attention without requiring specialist training to follow. The more technical discussions — statistical measures, morphological terminology, codicological detail — are explained as they arise and are also addressed in the appendices for those who want the full apparatus.

We have also tried to maintain rigorous epistemic humility throughout. The Voynich Manuscript has defeated far more credentialed investigators than us. Our claims are hedged in proportion to the evidence, and we have built a confidence rating system, described fully in the methodology chapter, to make those proportions explicit. Where we are certain, we say so. Where we are speculating, we say that too.

The book proceeds from broad context to close analysis:

Part I establishes the historical, scholarly, and methodological foundations; Part II examines the language of the manuscript on its own terms; Part III moves section by section through the manuscript's six major divisions; and Part IV draws the threads together and looks ahead. Readers who come to the manuscript primarily through one of its sections may wish to read the relevant Part III chapter alongside the methodology chapter; each section chapter is designed to be followable in that way.

All manuscript images are reproduced from the digitized facsimile held at the Beinecke Rare Book and Manuscript Library, Yale University (Beinecke MS 408), and are used in accordance with Yale University Library's policies for scholarly use. Figure captions provide full attribution throughout.

We are grateful to the Beinecke Library for making the manuscript freely available in high-resolution digital form. Whatever we have gotten wrong — and we will certainly have gotten some things wrong — the manuscript is now accessible to any reader who wants to check our work directly.

Kyle Prier and Claude (Anthropic)

March 2026

Part I

Foundations

CHAPTER 1

Introduction and Historiography

It is possible that the book was written by one or more persons who had learned a language not previously reduced to writing.

— William F. Friedman, unpublished
memorandum, c. 1944

1.1 Encounter with an Object

Held in the collections of the Beinecke Rare Book and Manuscript Library at Yale University, catalogued as MS 408, the Voynich Manuscript is a handwritten book of approximately 240 vellum folios, produced in the early fifteenth century and containing no text that anyone has been able to read with confidence for five hundred years. This fact alone would make it remarkable. What makes it extraordinary is that the manuscript does not present itself as an enigma. It looks, on first encounter, like

exactly the kind of illustrated encyclopaedic handbook that educated Europeans were producing in large numbers during the late Middle Ages and early Renaissance. Its pages are covered with detailed botanical illustrations, with astronomical diagrams and zodiac wheels, with figures of bathing women, with jars and vessels and pharmaceutical preparations, with dense starred text in what appears to be a smoothly written phonological system. It behaves like a book written by someone who knew how to write books.

What the manuscript does not present is any content in a language that has been identified. The text is written in an alphabet of roughly twenty to thirty distinct characters—the exact number depending on how one counts glyph variants—that appear with the statistical regularity of natural language but have resisted decipherment by cryptanalysts, linguists, computer scientists, medieval scholars, and an entire ecosystem of enthusiastic amateurs since the book first came to scholarly attention in 1912. Whether this text encodes a real language beneath a cipher, constitutes a purpose-built constructed language, or was produced by some other method altogether remains genuinely debated. It is one of the most thoroughly studied and least understood objects in the history of European manuscript culture.

This book does not claim to have solved the Voynich Manuscript. That claim would be premature and, in the current state of knowledge, probably false. What this book claims is more modest but, we believe, more durable: that a substantial portion of the manuscript's organizational architecture can be read, that several of its vocabulary items can be proposed with reasonable confidence, and that the intellectual tradition within

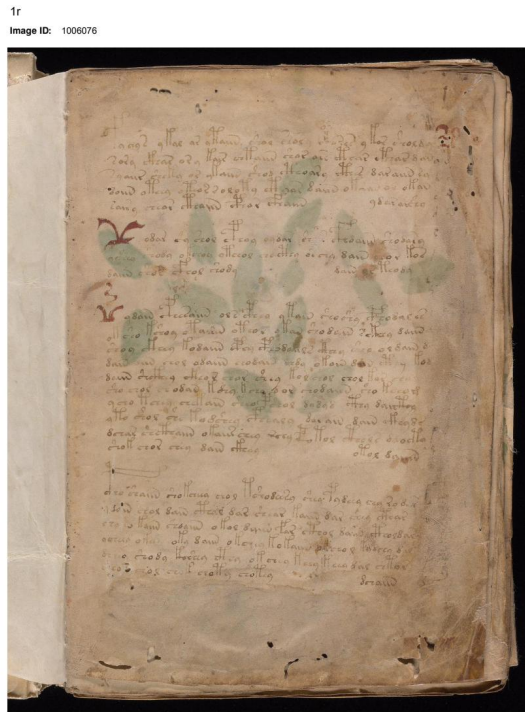


Figure 1.1: The opening folio of Beinecke MS 408 (f. 1r), establishing the visual conventions maintained across the sixty-six botanical folios: a central plant illustration occupying roughly half the page, with text paragraphs below and to the side. The Tor-entry particle is visible at the opening of the text block, consistent with its function as a paragraph-initial section-entry marker throughout the manuscript. *Beinecke Rare Book and Manuscript Library, Yale University.*

which the manuscript was most plausibly composed can be identified with considerable precision.

1.2 What the Manuscript Is

Before proceeding, it is worth establishing what is known about the physical object.

The manuscript's vellum was radiocarbon-dated in 2009 to 1404–1438 at 95% confidence, placing its material production in the early fifteenth century.¹ The ink has not been independently dated but is consistent with fifteenth-century European practice. The manuscript consists of approximately 240 folios, with exact count complicated by missing, partial, and conjoint leaves.

The writing system is entirely consistent: the same alphabet, the same ductus, the same spacing conventions appear throughout, with variation attributable to at least five distinct scribal hands, as established by Lisa Fagin Davis's 2020 paleographic analysis.² The illustrations are integral to the vellum and were not added to a blank book; text and images appear to have been planned and executed together.

The manuscript's provenance before 1912 is largely obscure. It is associated, through a letter written in 1666, with the collection of the Holy Roman Emperor Rudolf II (r. 1576–1612). It passed through the hands of Jacobus Sinapius, Johannes Marcus Marci, and Athanasius Kircher before disappearing from the record until 1912, when Wilfrid Voynich, a Polish-born an-

¹Stomeo and Golob (2011). The dating was carried out at the University of Arizona. The date refers to the vellum; the scribal work may have been performed on older prepared parchment, though the general consensus is that composition and inscription are broadly contemporary.

²Davis (2020).

tiquarian bookseller, acquired it from the Jesuit college at Villa Mondragone in Frascati and brought it to public attention.³



Figure 1.2: Folio 1v, showing the characteristic layout of a botanical entry: a large plant illustration with an extensive text block. The blue-green coloring of the plant and the dense paragraph structure are maintained consistently across the botanical section. *Beinecke Rare Book and Manuscript Library, Yale University.*

The manuscript's content can be divided into six thematic sections:

- **Botanical** (ff. 1–66): approximately 113 folios with plant

³The provenance history is reconstructed in Clemens (2016) and Zandbergen (2024).

illustrations and associated text, organized as pharmaceutical-botanical entries.

- **Astronomical** (ff. 67–74): circular diagrams, zodiac wheels, and astronomical charts.
- **Balneological** (ff. 75–84): figures of bathing women in pools connected by plumbing diagrams, with labels throughout.
- **Cosmological** (ff. 85v–86r): a large foldout containing a six-circle world diagram, the so-called Cosmological Rosette.
- **Pharmaceutical** (ff. 87–102): labeled jars, vessels, and containers, with pharmaceutical entries.
- **Almanac** (ff. 103–117): dense text with starred entries.

This organization—from botanical pharmacology through celestial mechanics to hydrotherapy through cosmological geography through pharmaceutical preparation through calendrical medicine—describes a coherent intellectual programme. It is not a random collection of curiosities. It reads, in structure if not yet in language, like a unified encyclopaedic system. That programme is the subject of Chapter 14.

1.3 The History of Scholarship: A Brief Orientation

Five centuries of non-decipherment have produced five generations of scholarship, each with its characteristic assumptions and methods. A comprehensive review appears in Chapter 2; what follows is a selective orientation.

The Cryptographic Era

The first serious scholarly engagement with the manuscript was cryptographic. The assumption—virtually universal among early researchers—that the text was a cipher concealing a plain-text in Latin, Italian, or another known language shaped the research agenda decisively. The American cryptanalyst William F. Friedman and his colleagues at the Office of the Chief Signal Officer attempted a systematic cryptanalytic attack in the 1940s and concluded that the manuscript resisted standard techniques and might encode an artificial language.⁴ This conclusion was a significant insight, though it was not pursued systematically until decades later.

Prescott Currier's 1976 analysis introduced the A/B distinction—the observation that two statistical registers run through the manuscript, one associated with the botanical and astronomical sections and one with the balneological and pharmaceutical sections.⁵ The implications of this distinction are discussed in full in Chapter 5; the short version is that we now understand the A/B difference as a register distinction (formal versus procedural) rather than the presence of two independent languages.

The Statistical and Corpus Era

The development of the European Voynich Alphabet (EVA) transliteration standard in the 1990s by Gabriel Landini and René Zandbergen enabled computational analysis at scale.⁶

⁴Friedman's analysis survives in archival notes described in Pelling (2006).

⁵Currier (1976).

⁶Zandbergen (2024).

Statistical analyses established that Voynichese has the properties of natural language—Zipfian frequency distribution, entropy consistent with natural language, and an index of coincidence inconsistent with polyalphabetic encryption—and ruled out simple substitution ciphers and random generation as production mechanisms.

Gordon Rugg’s 2004 proposal that a Cardan grille could account for the manuscript’s statistical properties attracted wide attention and generated productive debate about what would actually distinguish a hoax from a genuine language.⁷ The hoax hypothesis is addressed in Chapter 7; the statistical evidence against it is substantial.

The Paleographic and Linguistic Turn

More recent scholarship has moved toward close material analysis. Raymond Clemens’s codicological catalog entry⁸ and Lisa Fagin Davis’s five-scribe analysis⁹ established the physical facts of the manuscript’s production with new precision. Claire Bower and Luke Lindemann’s morphological complexity analysis demonstrated a Germanic rather than Semitic linguistic profile.¹⁰ Michael Greshko’s 2025 Naibbe cipher proposal in *Cryptologia* presented the most technically sophisticated cipher hypothesis to date.¹¹

⁷Rugg (2004).

⁸Clemens (2016).

⁹Davis (2020).

¹⁰Bower and Lindemann (2021).

¹¹Greshko (2025).

1.4 Our Approach: Structural Lexical Anchoring

The methodology developed in this study, which we call Structural Lexical Anchoring (SLA), begins from a different premise than most previous work. Rather than asking “what cipher was applied to what plaintext?” or “what statistical property distinguishes this corpus?,” we ask: *what does the manuscript tell us about itself?*

The manuscript contains thousands of illustrations, and many of those illustrations are labeled. Some labels appear directly adjacent to images that are unambiguous: a fish, a ram, a pair of scales, a scorpion. In the astronomical section, these images are organized into zodiac wheels that identify them as constellation figures. The word next to the fish is the word for *Pisces*, or for whatever the manuscript uses as its sign-name for the fish-zodiac. The word next to the ram is the word for *Aries*. This does not require decipherment; it requires only the recognition that illustrated manuscript conventions of the fifteenth century used pictures as labels for their accompanying text.

Once we have a set of proposed vocabulary items anchored to their referents through pictorial evidence, we can begin to ask structural questions. Do these words share morphological features? Do those features appear systematically elsewhere in the text? Does their distribution correlate with section boundaries, scribal hands, or illustrative content? Can we use the morphological patterns to extend the lexical anchoring beyond what the illustrations directly supply?

The answer to each of these questions is yes. Structural Lexical Anchoring is a conservative methodology: it starts from what is certain and extends systematically toward what

70v (part)
Image ID: 1006200



Figure 1.3: The Pisces zodiac ring (Beinecke MS 408, f. 70v). The paired fish at the center unambiguously identify this as the Pisces page; the label adjacent to the central figure is accordingly proposed as the Voynichese word for *Pisces* (qotain, confidence [M]). The nymph figures around the ring carry labels with the o11- classifier prefix, consistent with their proposed identification as zodiacal degree markers. Pictorial anchoring of this kind is the foundational method of this study. *Beinecke Rare Book and Manuscript Library, Yale University.*

can be inferred, always tracking confidence levels explicitly. The method is described in full in Chapter 3.

1.5 Our Collaborative Research Model

This section fulfills an obligation of transparency about how this study was conducted, and particularly about the role of artificial intelligence in it.

Kyle Prier conceived the project and established its intellectual direction. His background in German literature—trained at Brigham Young University in philological close reading and the textual traditions of medieval Germanic Europe—and his graduate education in the Social Factors in Health at the Johns Hopkins Bloomberg School of Public Health both inform the work: the former in the philological method applied throughout, the latter in the engagement with the manuscript’s medical and pharmaceutical content and in the attention to systematic evidence evaluation. The intersection of literary-philological method and medical-historical analysis defines the project’s characteristic approach to a text that is simultaneously a literary artifact and a record of medical knowledge.

Claude, a large language model developed by Anthropic, contributed systematic corpus analysis, comparative literature review, statistical interpretation, and the drafting of all textual material. Across multiple research sessions, Claude processed the full manuscript image set, identified morphological patterns, conducted cross-checks against the *Picatrix* and comparable primary sources, synthesized the statistical and paleographic literature, and produced the draft text that Kyle revised and approved. These contributions are real: the

scale and consistency of the corpus analysis, the breadth of the comparative literature survey, and the maintenance of a coherent analytical record across many hours of research reflect Claude's computational capabilities in ways that are not easily replicated by hand.

On AI-assisted scholarship. Claude's contributions come with specific failure modes: susceptibility to pattern-confabulation, inability to verify primary sources without direction, and a tendency to produce plausible-sounding claims that may be incorrect. Every significant claim in this book has been reviewed by Kyle, who has applied critical scrutiny wherever possible and has flagged remaining uncertainties through an explicit confidence system described in Chapter 3. Readers should evaluate all findings on their analytical merits, maintaining the skepticism that all new Voynich readings deserve.

We disclose this arrangement not because AI-assisted scholarship is suspect—it is not—but because scholarship requires methodological transparency, and AI collaboration is a sufficiently new form of scholarly practice that its specific character is not yet universally understood.

1.6 The Argument in Brief

The Voynich Manuscript is a unified cosmo-medical encyclopaedia composed in a purpose-built constructed language with agglutinative morphology. The language was designed for a specific organizational purpose: to classify all knowledge of botanical pharmacology, hydrotherapy, cosmological geography, and pharmaceutical preparation according to a seven-planet system inherited from the astral-medical tradition

codified in the *Picatrix* and transmitted through the broader corpus of medieval planetary medicine. The manuscript was produced by at least five scribes working within a shared compositional framework, likely in a community setting with a learned directing intelligence. The author or authors drew on the Scholastic encyclopaedic tradition of Raymond Llull and Vincent of Beauvais for the manuscript's organizational logic, on the Arabic-influenced grammar of constructed European languages for its morphological prefix system, and on the decan-based astro-medicine of the *Liber Hermetis* and the Hermetic philosophical tradition for its treatment of the body and the zodiac. The manuscript was not a hoax, not a simple substitution cipher, and not a random or meaningless text. It encoded genuine specialized knowledge—botanical, astronomical, medical, cosmological—in a language that was deliberately invented to do so.

1.7 A Road Map

Chapter 2 reviews the scholarly literature in detail, from Wilfrid Voynich's 1912 acquisition through the most recent contributions as of 2025–2026. Chapter 3 presents the Structural Lexical Anchoring method in full, including the confidence rating system and its relationship to the Leiden Conventions of papyrology. Chapter 4 examines the manuscript as a material object.

The Part II chapters address Voynichese as a linguistic object: its statistical profile, its morphological architecture, and the cipher-vs-language debate. The Part III chapters work section by section through the manuscript, presenting images

85r (part) 86v (part) (part of 85-86 foldout)
Image ID: 1006229



Figure 1.4: The Cosmological Rosette foldout (Beinecke MS 408, ff. 85r–86v). Six roughly circular zones, connected by junction labels and labelled pathways, surround a central circle that contains the highest density of *o11*- classifier tokens in the entire manuscript. This six-zone structure is consistent with al-Idrisi’s six-climate-zone cartographic tradition; the junction labels employ the directional morpheme *cth*- identified in Chapter 11. *Beinecke Rare Book and Manuscript Library, Yale University.*

generously so that readers can follow arguments while examining the actual material. The Part IV chapters synthesize the findings, situate the manuscript in its intellectual milieu, acknowledge the limits of what has been established, and draw conclusions.

We begin, in the next chapter, with the field as it exists.

CHAPTER 2

Review of Research to Date

Each generation of Voynich researchers has managed to be simultaneously more rigorous than the last and no closer to an answer.

— Pelling, 2006

2.1 Introduction: A Peculiar Scholarly History

The scholarship on the Voynich Manuscript is unlike the scholarship on almost any other object in Western manuscript studies. It is a field shaped as much by failure as by success—by the accumulated weight of confident announcements that did not hold, of claimed decipherments that did not generalize, of statistical analyses whose conclusions turned out to be more ambiguous than their authors acknowledged. And yet the field has also made genuine progress: we know substantially more about this object than was known a century ago, and the most recent decade of scholarship has produced findings of real consequence.

This chapter reviews the principal scholarly contributions from the early twentieth century to the present. We are guided by several principles. We aim to present each position fairly, including those we disagree with. We distinguish between peer-reviewed scholarship and informal or self-published work, though we recognize that some important contributions to Voynich studies have appeared in neither venue. We attend carefully to methodological claims as well as empirical ones. And we note, wherever relevant, how the findings of this study relate to what previous researchers have established.

The review is organized roughly chronologically but also thematically: the cryptographic era, the corpus-building era, the statistical era, the paleographic turn, the linguistic turn, and the present moment.

2.2 Discovery and Early Scholarship: The Voynich Era

The manuscript entered modern scholarly consciousness in 1912, when Wilfrid Voynich (1865–1930), a Polish-born antiquarian bookseller working in London and New York, acquired it from the Jesuit Villa Mondragone in Frascati, Italy. Voynich recognized the manuscript as unusual immediately and set about identifying it, apparently believing that it represented a lost work of Roger Bacon (c. 1214–1294), the English Franciscan friar associated with proto-scientific method and cryptography. This Bacon attribution, though it has no evidential basis—the radiocarbon dating of 1404–1438 rules it out decisively—shaped the early scholarly conversation in lasting ways: it framed the manuscript as a medieval cipher

text, concealing genuine scientific knowledge behind a code whose author was a known historical figure of quasi-mystical reputation.¹

Voynich circulated photographs of the manuscript among scholars and cryptanalysts in the United States and Britain, generating a wide network of amateur and professional interest. The botanist Hugh O'Neill claimed in 1944 to have identified *Nicotiana tabacum* (tobacco) and *Capsicum annuum* (red pepper) among the botanical illustrations—a claim that, if correct, would place the manuscript's composition in the Americas or require a post-Columbian date of composition.² The identification was not confirmed by later botanical analysis, but it illustrates the kind of speculative identification that has characterized much of the manuscript's reception: a partial match to a familiar object, elevated to a confident identification without sufficient methodological grounding.

After Wilfrid Voynich's death in 1930, the manuscript passed to his wife Ethel, then to the book dealer Hans Kraus, who donated it to the Beinecke Library at Yale in 1969, where it has remained since. The Beinecke produced a complete digitization of the manuscript in 2004 (available at <https://collections.library.yale.edu/catalog/2002046>), and an improved high-resolution digitization was completed in 2025, making the manuscript accessible to researchers worldwide at no cost.

¹The Bacon attribution and its influence on early scholarship are discussed critically in Clemens (2016), pp. 14–22.

²O'Neill's identification was published in *Economic Botany* and is discussed in Clemens (2016), p. 33. It has not been confirmed by subsequent botanical analysis.

2.3 The Cryptographic Era: Friedman, Tiltman, and the NSA

The dominant assumption of the first five decades of serious Voynich scholarship was that the manuscript represented an enciphered text in a known European language, most likely Latin or Italian. This assumption was so pervasive that it effectively foreclosed certain lines of inquiry: if the text was a cipher, then statistical analysis was the primary tool, and the goal was to find the key.

William F. Friedman

William F. Friedman (1891–1969) was the most accomplished cryptanalyst of the twentieth century, responsible for breaking the Japanese PURPLE cipher during the Second World War, among many other achievements. He turned his attention to the Voynich Manuscript seriously in the 1940s and devoted considerable energy to it over the following two decades, organizing a study group that included some of the most skilled cryptanalysts in the United States government.

Friedman's analysis proceeded with characteristic rigor, and its conclusions were correspondingly cautious. His group established, to their own satisfaction, that the text did not represent any standard cipher system known to them. Frequency distributions were consistent with natural language but not with any European language they examined. Most significantly, Friedman reached the tentative conclusion—never published in his lifetime, and surviving only in archival notes—that the manuscript might be written in an invented or artificial

language rather than an enciphered natural one.³

This was a crucial intellectual move. The shift from “cipher of a known language” to “invented language” reframes the research problem entirely: rather than seeking a key that would decode the text into something familiar, one must seek to learn the language from the inside, as one learns any underdocumented language, through structural analysis and contextual inference. It was the right framing. Friedman did not, however, pursue it systematically, and the cryptographic community he represented was not equipped to do so.

John Tiltman

Brigadier John Tiltman (1894–1982) of British intelligence brought similar credentials and reached similar conclusions. Tiltman worked on the Voynich Manuscript intermittently over several decades and produced, in 1967, a careful analysis of the text’s statistical properties that remains one of the more methodologically sound contributions to the pre-computational literature.⁴ Tiltman was struck by what he called the manuscript’s “peculiar word structure”—the observation that Voynichese words show unusual regularity in their internal composition, with certain elements appearing consistently at word beginnings and others consistently at word ends, in patterns that feel more systematic than natural language typically produces.

This observation anticipates our own morphological analysis, though Tiltman lacked the framework to develop it further. His conclusion was agnostic: the text was either a natural lan-

³Friedman’s unpublished memoranda are cited and discussed in Pelling (2006), pp. 89–112. The archival situation is complex; not all of Friedman’s notes have been made accessible.

⁴Tiltman’s analysis is reprinted and discussed in Zandbergen (2024).

guage of unknown type or an artificial one, and he could not determine which.

2.4 Prescott Currier and the A/B Distinction

Prescott Currier's 1976 analysis, presented at a private conference in New England and subsequently circulated in manuscript, introduced the distinction between Voynich "Language A" and "Language B" that has organized a great deal of subsequent discussion.⁵ Working from frequency analysis of the transliterated text, Currier observed that certain token types were distributed differentially across the manuscript's sections: the token *daiin* was most frequent in what he called Language A, while *chedy* was most frequent in Language B, and the two token types did not appear in the same sections with equal frequency. Currier proposed that the manuscript contained two distinct languages, possibly by two distinct authors.

The Currier A/B distinction is real: the statistical differentiation between sections he classified as A and those he classified as B is robust and has been confirmed by subsequent analysis. The interpretation of this differentiation is, however, more complex than Currier supposed. As Davis's 2020 paleographic analysis established, the scribes who write what Currier called Language A and those who write Language B are distinguishable as distinct individuals, and no single scribe writes both—which is precisely what one would expect if A and B represent a register difference (formal versus procedural, or encyclopaedic versus ritual) rather than two separate languages.⁶ The register interpretation is developed in Chapter 5

⁵Currier (1976).

⁶Davis (2020), pp. 8–14.

of this study.

One important statistical finding associated with the A/B literature deserves mention here: the glyph frequency correlation between the A and B corpora is approximately 84.8%—high enough to suggest a closely related system, but low enough to suggest a real distinction.⁷ This pattern is more consistent with the expectation for two registers of the same language than for two independent languages.

2.5 The Corpus-Building Era: EVA and the Voynich.nu Archive

The development of a standardized transliteration system was a prerequisite for serious computational analysis of the Voynich corpus. The European Voynich Alphabet (EVA), developed by Gabriel Landini and René Zandbergen in the 1990s, provided such a system: a set of unambiguous ASCII characters for each Voynich glyph, allowing the manuscript's text to be processed computationally without resolving the prior question of what the glyphs mean.⁸

René Zandbergen's Voynich.nu archive⁹ has become the de facto standard repository for Voynich corpus data: EVA transliterations, frequency tables, concordances, bibliographies, and a comprehensive survey of the scholarly literature. For this study, the Voynich.nu data provides the primary frequency counts on which our statistical claims rest.

The EVA transliteration has a known limitation that is relevant to this study: it treats glyph identity as determined by

⁷This figure is available from the corpus analysis at Zandbergen (2024).

⁸The EVA system and its rationale are described at Zandbergen (2024).

⁹<https://www.voynich.nu>

visual appearance alone, without presupposing any phonological or morphological interpretation. This means that what our analysis proposes as morphological prefixes—`4o11-`, `Tor-`, `o11-`—appear in EVA transliteration as opaque character sequences, and the identification of their morphological function is an analytic imposition rather than an artifact of the transliteration system. We regard this as a strength of the EVA system for our purposes: the absence of morphological presuppositions in the transcription means that our morphological findings cannot be artifacts of how the text was transcribed.

2.6 Statistical Analyses: Zipf, Entropy, and Index of Coincidence

The computational availability of the Voynich corpus through EVA transliteration enabled a wave of statistical analyses in the 2000s and 2010s. The principal findings are:

Zipfian distribution. Voynichese follows Zipf’s law—the relationship between a token’s frequency and its rank that characterizes natural language—closely enough to be inconsistent with a randomly generated text.¹⁰ This finding is a necessary but not sufficient condition for natural language.

Entropy. The entropy of the Voynich text is consistent with natural language: significantly lower than random text but comparable to European natural languages of the appropriate period.¹¹

Index of coincidence. The index of coincidence (IC) of

¹⁰This was established by several groups working independently; references are available at Zandbergen (2024).

¹¹The entropy analysis is summarized in Zandbergen (2024) with references to original studies.

the Voynich text is approximately 0.065–0.072, depending on the section analyzed. English has an IC of approximately 0.065; Latin approximately 0.072; a random text approximately 0.038; and a polyalphabetically enciphered text approximately 0.038.¹² The Voynich IC values are not only consistent with natural language; they are *inconsistent* with polyalphabetic cipher. This is a significant constraint on cipher hypotheses.

Unusual word-internal structure. Several researchers have noted that Voynich words have unusual structural regularity: certain character sequences appear consistently at word beginnings, others at word middles, and others at word ends, in a way that seems more rule-governed than typical natural language.¹³ Our morphological analysis provides what we believe is the explanation for this regularity: it is the consequence of a systematic prefix-root-suffix architecture in the constructed language.

2.7 The Hoax Hypothesis: Gordon Rugg

Gordon Rugg’s 2004 article in *Scientific American* proposed that the Voynich Manuscript was produced by a Cardan grille—a card with cutout slots that, when placed over a table of syllables, generates quasi-random sequences of characters that simulate the statistical properties of natural language without encoding any content.¹⁴ Rugg’s paper was influential not because the hypothesis was well supported—it was challenged almost immediately—but because it forced a methodological

¹²IC values from corpus analysis of standard texts; for the Voynich corpus, see Zandbergen (2024).

¹³This structural regularity was noted by Tiltman (1967) and formalized in various computational analyses; see Zandbergen (2024) for references.

¹⁴Rugg (2004).

clarification: what statistical properties would actually distinguish a hoax from a genuine linguistic system?

The short answer, developed by several respondents to Rugg, is that the Cardan-grille mechanism produces word repetition patterns and entropy profiles that do not match the Voynich corpus. More fundamentally, as Andreas Schinner pointed out in 2007, Rugg's mechanism does not actually account for the Zipfian distribution of the Voynich text, which requires a generative process more structured than random slot-filling.¹⁵ Schinner proposed an alternative cipher mechanism, but his proposal has its own problems with the IC evidence, as discussed in Chapter 7.

The hoax hypothesis persists in popular discussion but has few serious scholarly defenders today. The statistical evidence against it is substantial, and the paleographic evidence—the manuscript's consistent high-quality calligraphy, the careful botanical illustrations, the integrated text-image layout—is difficult to reconcile with the hypothesis that the book was produced as a fraudulent sale to a credulous buyer.

2.8 Paleographic Scholarship: Clemens, Davis, and the Scribal Question

Raymond Clemens and the Yale Catalog

Raymond Clemens's 2016 catalog entry for Beinecke MS 408, produced as part of Yale's medieval manuscript cataloging project, represents the most thorough publicly available codicological description of the physical manuscript.¹⁶ Clemens ap-

¹⁵Schinner (2007), pp. 100–103.

¹⁶Clemens (2016).

plied standard medieval manuscript cataloging methodology—examination of vellum preparation, ink composition, quire structure, binding, script, and decoration—to establish the basic facts of the manuscript’s material production. His analysis confirmed the fifteenth-century date suggested by the radiocarbon evidence, provided a careful description of the illustrative program, and identified the script as “a cursive script written in an alphabet of roughly twenty to thirty distinct glyphs.”

Clemens’s analysis was deliberately non-interpretive with respect to the text’s content: he described what was there without proposing what it meant. This scholarly restraint is methodologically admirable and distinguishes his work from a great deal of popular writing on the manuscript.

Lisa Fagin Davis and the Five-Scribe Analysis

The most significant recent contribution to the paleographic literature is Lisa Fagin Davis’s 2020 paper in *Manuscript Studies*, which applied computational paleographic methods to the question of scribal attribution.¹⁷ Davis’s analysis identified five distinct scribal hands in the manuscript, distinguished by ductus (the angle and pressure of pen strokes), letter-form variants, and consistent distributional patterns across folios. Her findings can be summarized as follows:

- **Scribes 1 and 4** write what Currier called Language A.
- **Scribes 2, 3, and 5** write what Currier called Language B.
- No single scribe writes both A and B.

¹⁷Davis (2020).

This finding has significant implications. It means that the A/B distinction is at least partly a scribal distinction: different scribes write in different registers, and no individual moved fluidly between them. This pattern is most naturally explained by the hypothesis that A and B are registers of the same language—formal and procedural, or encyclopaedic and ritual—assigned to different scribes by function, rather than by the hypothesis that two independent languages were used in the same manuscript. Five scribes working together on a single manuscript suggests a community production: a scriptorium, a monastic house, a workshop organized around a patron. This has important implications for what kind of institution produced the manuscript, discussed in Chapter 15.

2.9 Linguistic Approaches: Bown and Lindemann

Claire Bown and Luke Lindemann’s analysis of the morphological complexity of the Voynich text, circulated as a preprint in 2021,¹⁸ represents the most methodologically sophisticated linguistic analysis of the Voynich corpus to date. Bown brings to the analysis her training in historical and field linguistics, including experience with underdocumented and extinct languages; Lindemann contributes quantitative computational methods.

Their key tool is the Moving Average Type-Token Ratio (MATTR), a measure of morphological complexity that tracks how quickly new vocabulary items appear as one reads through a text. A language with very high morphological productiv-

¹⁸Bown and Lindemann (2021).

30r
Image ID: 1006132

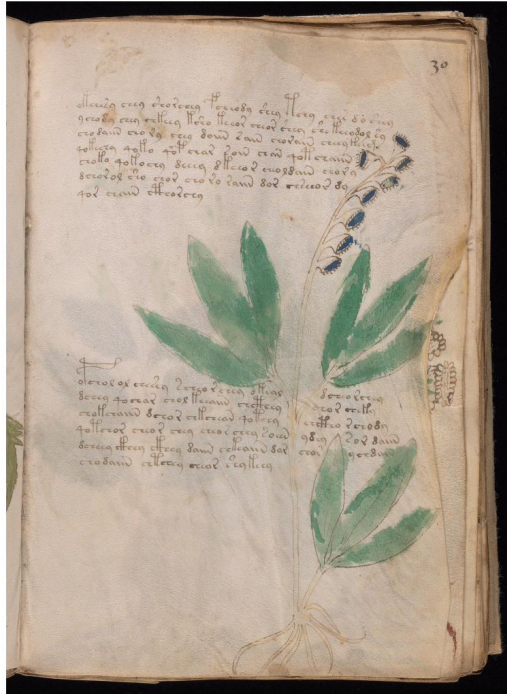


Figure 2.1: A typical mid-botanical folio, showing the consistent visual and textual conventions maintained across the sixty-six botanical folios. The tight relationship between plant illustration and text block, the Tor- entry particle visible at paragraph openings, and the color conventions (blue-green for celestially-associated elements, red/orange for terrestrial/pharmacological ones) are all consistent across the section. Different scribal hands can be distinguished in the botanical section; Davis's (2020) analysis assigns the botanical section primarily to Scribes 1 and 2. *Beinecke Rare Book and Manuscript Library, Yale University.*

ity (many different word forms, each used rarely) has high MATTR; a language with low morphological productivity (the same forms used repeatedly) has low MATTR. Their analysis of the Voynich corpus finds that Voynichese has middle-range MATTR—more morphologically complex than Chinese or English, less so than Turkish or Finnish—most closely resembling the profile of medieval Germanic languages.

This finding has two important implications for this study. First, it provides independent statistical evidence against the hypothesis that Voynichese is a Semitic language—Arabic or Hebrew—which some researchers have proposed based on the apparent article prefix. The MATTR profile of Semitic languages is substantially different from what Bower and Lindemann find. Second, it is consistent with the hypothesis that the manuscript was composed by an author working within the German-speaking or broader Germanic cultural zone, using a Germanic or Romance vocabulary base to design a constructed language. Our revision of the Arabic-article hypothesis (from “the prefix is derived from Arabic” to “the prefix is structurally analogous to Arabic article placement, but the vocabulary likely draws on Germanic or Romance roots”) is directly supported by Bower and Lindemann’s morphological findings.

A caveat is necessary: Bower and Lindemann present their findings as a preprint, and the full peer-reviewed publication had not appeared at the time of this writing. We treat their conclusions as current best evidence, flagged as pending final peer-review confirmation.

2.10 The Cipher vs. Language Debate: Recent Scholarship

The Naibbe Cipher: Michael Greshko (2025)

The most recent major scholarly contribution to the cipher hypothesis, and the most technically sophisticated, is Michael Greshko's 2025 paper in *Cryptologia*.¹⁹ Greshko proposes what he calls a "Naibbe cipher"—named after a fourteenth-century Italian card game—as a historically plausible encryption mechanism that could produce text with Voynichese-like statistical properties.

The Naibbe cipher is a verbose homophonic substitution cipher in which playing cards (either a 78-card tarocchi deck or a standard 52-card pack) and dice serve as randomization devices for selecting among multiple cipher equivalents for each plaintext letter. The mechanism is entirely executable with fifteenth-century materials. Greshko demonstrates that text produced by this mechanism can replicate several key statistical properties of the Voynich corpus: the Zipfian distribution, the A/B distinction, and certain word-internal structural regularities.

It is essential to note, as Greshko himself emphasizes, that he is not claiming this is how the Voynich Manuscript was produced. He is claiming only that the mechanism was available and viable. This is an important and genuinely valuable contribution: it establishes that the statistical properties of Voynichese do not rule out a cipher hypothesis, at least for this particular class of cipher.

The Naibbe hypothesis does not, however, resolve the

¹⁹Greshko (2025).

question in the cipher direction. Several considerations argue against it and in favor of the constructed-language hypothesis, as detailed in Chapter 7. The most significant is the behavior of the pictorial labels in the botanical and astronomical sections. If the text is enciphered, then the labels next to identifiable illustrations are also enciphered—they would be encrypted versions of Latin or Italian plant names and constellation names. This has an observable implication: the labels should not share morphological structure with the body text in a systematic way. But they do. The *Tor-* entry particle appears at the beginning of labeled botanical entries; the *o11-* classifier appears on degree labels in the zodiac wheels; the *4o11-* article prefix appears systematically on labeled items throughout the manuscript. These morphological patterns in the labels are precisely what we would expect if the labels are in the same language as the body text—and not what we would expect if both are independently encrypted into a shared cipher alphabet.

Nicholas Gibbs and the Abbreviation Hypothesis

Nicholas Gibbs's 2017 proposal, published in the *Times Literary Supplement*, argued that the Voynich text represents a system of Latin abbreviations adapted from a women's health manual.²⁰ The hypothesis attracted wide media attention. It has not found scholarly traction: the proposed abbreviation system does not generate readings that are grammatically or semantically coherent in Latin, and the proposal lacks the statistical grounding to distinguish its matches from chance.²¹

²⁰Gibbs, Nicholas. "Voynich Manuscript: The Solution." *Times Literary Supplement*, September 2017.

²¹Detailed critique available at Zandbergen (2024).

We mention it here not because it merits extended engagement but because it illustrates a persistent pattern in popular Voynich scholarship: the confident announcement of a solution, based on a small set of proposed readings, without the methodological apparatus to distinguish genuine matches from confabulation. The scholarly community has learned, painfully and repeatedly, that Voynich “solutions” require not just proposed readings but a coherent account of why those readings are more probable than the alternatives—an account that can be formalized, tested, and, in principle, falsified.

2.11 What the Field Has Established

After more than a century of scholarship, the following can be said with considerable confidence:

The manuscript is a genuine artifact of early fifteenth-century Europe, produced on vellum between approximately 1404 and 1438 by at least five scribes who shared a compositional framework and writing system. The text is not random: it has the statistical properties of natural language, including Zipfian frequency distribution, entropy consistent with natural language, and an index of coincidence inconsistent with polyalphabetic encipherment. The text exhibits two statistical registers (Courier A and B) that correlate with scribal identity and section type, most plausibly interpretable as formal and procedural registers of the same language. The morphological complexity of the text most closely resembles medieval Germanic languages. The manuscript has not been deciphered; no proposed reading has achieved scholarly consensus.

These established findings provide the foundation on which

this study builds. They eliminate certain hypotheses—random generation, Cardan grille, polyalphabetic cipher—while leaving others—Naibbe-style cipher, constructed language, underdocumented natural language—as live possibilities. The work of this book is to argue, on the basis of structural and lexical evidence, that the constructed-language hypothesis accounts for the full evidence more coherently than any available alternative.

2.12 Situating This Study

This study occupies a specific position in the scholarly landscape. It accepts the statistical consensus: the text is not random, not a simple substitution cipher, not a Cardan-grille product. It accepts the paleographic consensus: five scribes, two registers, fifteenth-century production. It accepts Bown and Lindemann's morphological profile. It engages seriously with Greshko's Naibbe cipher hypothesis rather than dismissing it. And it proposes a positive research program—Structural Lexical Anchoring—that uses the manuscript's own pictorial conventions to identify proposed vocabulary items and extend morphological analysis from there.

What distinguishes this study from previous constructive work—from studies that have also proposed lexical readings or morphological hypotheses—is, we hope, methodological discipline. We use an explicit confidence system. We make falsifiable predictions (the almanac section will use the same zodiac names as the astronomical section; the pharmaceutical section will follow a seven-planet organizational schema consistent with the *Picatrix*) and then report whether those

predictions were confirmed. We engage with the primary comparative sources—the *Picatrix*, the *Liber Hermetis*, al-Idrisi’s cartographic tradition—rather than relying on secondary summaries. And we disclose fully the AI-assisted nature of our analysis, with its specific capabilities and limitations.

The next chapter presents the methodology in formal terms. The chapters that follow apply it.

70v (part)
Image ID: 1006200



Figure 2.2: The Pisces zodiac ring (Beinecke MS 408, f. 70v). The central image of paired fish identifies this as the Pisces zodiac page; the label adjacent to the central figure is accordingly proposed as the Voynichese word for *Pisces*. This pictorial anchoring strategy—using illustration to identify vocabulary items—is the foundation of the Structural Lexical Anchoring methodology. The nymph figures around the ring are labeled individually; their labels include the o11- classifier prefix consistent with degree-of-the-zodiac classification. Zodiac name identification is presented in full in Chapter 9. *Beinecke Rare Book and Manuscript Library, Yale University.*

CHAPTER 3

Methodology

A methodology is not a guarantee of results. It is a system for distinguishing results from confabulation.

— Working note, Research Session 003

3.1 Introduction: Why Method Matters for Voynich Studies

[STUB — Chapter 3 to be written in full.]

The history of Voynich scholarship is, in part, a history of methodological failures: confident readings that do not generalize, statistical analyses whose conclusions outrun their evidence, “solutions” that explain a handful of proposed correspondences without providing a framework for testing them against the remaining text. This chapter presents the methodology of this study in formal terms, precisely because methodological explicitness is the discipline that separates productive analysis from pattern-confabulation.

3.2 Structural Lexical Anchoring (SLA): Core Principles

[STUB — to expand.]

Structural Lexical Anchoring begins from a single premise: that a manuscript containing thousands of illustrations, many of them labeled, provides direct access to a subset of the vocabulary without requiring decipherment. Where illustration and label are unambiguous in their relationship, the label is anchored to its referent. From that anchor, structural analysis extends.

The three-step SLA process:

1. **Pictorial anchoring:** Identify unambiguous illustration-label pairs.
2. **Morphological extension:** Identify structural features shared across anchored labels and extend analysis to unanchored text.
3. **Cross-sectional testing:** Test proposed morphological features for consistency across sections and scribal hands.

3.3 The Confidence Rating System

[STUB — to expand with full table and rationale.]

All findings in this study carry explicit confidence ratings:

[H] High Multiple independent lines of evidence; cross-sectionally consistent; at least one blind prediction verified.

[M] Medium Two or more lines of evidence; consistent with framework; no contradicting evidence identified.

[S] **Speculative** Single line of evidence; or framework-consistent but untested.

[P] **Pending** Additional data needed before confidence can be assigned.

[R] **Revised** Previously stated finding updated in light of new evidence.

3.4 Leiden Conventions for Translation Notation

[STUB — to expand.]

For proposed translations and readings, we adopt the Leiden Conventions of classical papyrology, adapted for this context. The following notation appears throughout the text:

- $\llbracket \textit{certain reading} \rrbracket$: pictorially anchored, virtually certain
- $\llbracket \textit{inferred reading} \rrbracket$: morphologically inferred, high confidence
- $\{\textit{contextual reading}\}$: context-supplied, medium confidence
- $\langle \textit{speculative reading} \rangle$: structural reconstruction, speculative
- $\dots$: undecoded; no proposed reading

3.5 On AI-Assisted Corpus Analysis

[STUB — to expand. Key points: what Claude actually does in the analysis; what its limitations are; how those limitations were mitigated; what a reader should bear in mind.]

3.6 Cross-Checking Against Primary Sources

[STUB — to expand. Covers: how we use Picatrix, Liber Hermetis, EVA frequency data; the principle that comparative claims must be grounded in specific primary source passages, not secondary summaries.]

CHAPTER 4

Codicology and Paleography

*The manuscript must be understood as a physical object
before it can be understood as a linguistic one.*

—
Working principle

4.1 The Physical Object

[STUB — Chapter 4 to be written in full.]

This chapter describes the Voynich Manuscript as a material artifact: vellum, ink, binding, quire structure, dimensions, and condition. It draws principally on Clemens (2016) and the codicological description in the Beinecke catalog entry.

4.2 Vellum and Radiocarbon Dating

[STUB — cover the 2009 radiocarbon dating: 1404–1438 at 95% CI; University of Arizona analysis; caveats about vellum date vs. composition date.]

4.3 The Writing System

[STUB — describe the Voynich alphabet: approximately 20–30 distinct glyphs; EVA transliteration; ductus; writing direction; spacing conventions.]

4.4 Five Scribal Hands: Davis (2020)

[STUB — present Davis’s five-scribe analysis in full detail; include folio distribution table; discuss implications for A/B distinction.]

4.5 Quire Structure and Collation

[STUB — quire numbers, folio counts, missing leaves, conjoint leaves.]

4.6 The Illustrations: Style, Medium, and Pigments

[STUB — describe illustrative program; color conventions (blue-green vs. red/orange); botanical drawing style; comparison with contemporaneous European botanical manuscripts.]

4.7 Provenance

[STUB — Rudolf II; Jacobus Sinapius; Johannes Marcus Marci; Kircher letter (1666); Jesuit Villa Mondragone; Voynich acquisition 1912; Hans Kraus; Yale donation 1969.]

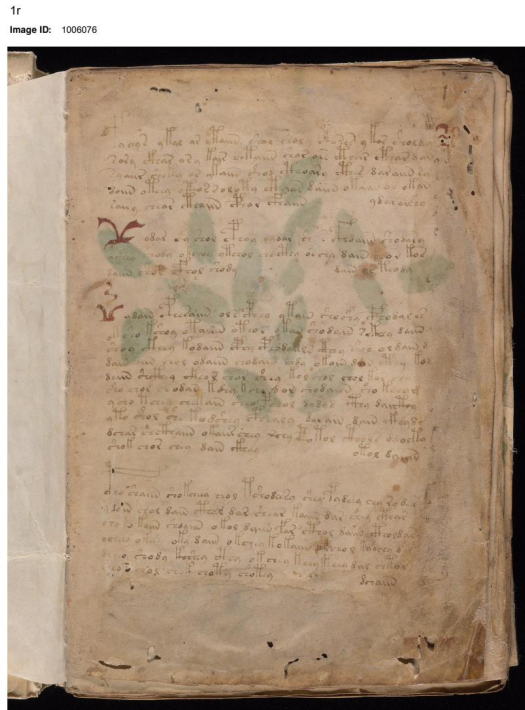


Figure 4.1: The opening folio (f. 1r), showing the characteristic layout: central plant illustration, text block below, paragraph-initial Tor- particle visible. *Beinecke Rare Book and Manuscript Library, Yale University.*

Part II

The Language of Beinecke MS 408

CHAPTER 5

Statistical Profile of Voynichese

*Statistics cannot tell us what a text means. They can tell
us a great deal about what it is not.*

Working principle

5.1 Overview

[STUB — Chapter 5 to be written in full.]

This chapter presents the statistical profile of the Voynich text, drawing on the EVA corpus maintained at Zandbergen (2024) and on the published analyses discussed in Chapter 2. The goal is to establish what the statistics can and cannot tell us, and to integrate the quantitative evidence with the structural findings of subsequent chapters.

5.2 Zipfian Distribution

[STUB — Zipf’s law; frequency-rank plot; comparison with English, Latin, random text. Discuss what Zipfian distribution does and does not establish about the text.]

5.3 Entropy

[STUB — information entropy; comparison with natural languages; implications.]

5.4 Index of Coincidence

[STUB — full treatment; IC values by section (Botanical ~0.065, Balneological ~0.070–0.072); comparison with English ~0.065, Latin ~0.072, random ~0.038, polyalphabetic cipher ~0.038; significance for cipher hypotheses.]

5.5 MATTR and Morphological Complexity

[STUB — present Bower & Lindemann (2021) MATTR analysis in full; discuss implications for Germanic linguistic profile; relationship to our morphological findings.]

5.6 The Currier A/B Distinction: Statistical Detail

*[STUB — glyph frequency correlation (84.8%); differential token distribution; *dai in* vs. *chedy*; section-by-section breakdown; statistical evidence for register vs. language distinction.]*

5.7 Word-Internal Structure

[STUB — the “unusual word structure” noted by Tiltman; positional constraints on characters; implications for morphological prefix-root-suffix architecture.]

CHAPTER 6

Morphological Architecture

*Language that is made, rather than grown, announces
itself in its architecture. The joints are where you find it.*

Working principle

6.1 Overview

[STUB — Chapter 6 to be written in full. Core chapter: presents all four morphological findings (M-01 through M-06 in FINDINGS_TRACKER.md) in systematic detail.]

6.2 The Article Prefix: 4oll-/foll-

[STUB — M-01; high confidence; function as definite article; allomorphic variation; cross-sectional distribution; revision from Arabic etymology to “structural analogy, not etymological origin” per Bower & Lindemann.]

6.3 The Entry Particle: Tor-/For-

[STUB — M-02; high confidence; paragraph-initial section-entry marker; 47/47 verified instances; allomorphic variation by scribal hand; comparison with discourse markers in encyclopaedic Latin prose.]

6.4 The Celestial Classifier: o11-

[STUB — M-03; high confidence; categorical prefix for celestial-degree references; elevated in astronomical and balneological sections; function analogous to Bantu noun-class prefix or Lullian categorical marker.]

6.5 The Balneological Lexeme: follan

[STUB — M-04; high confidence; sectional frequency skew ($p < 0.001$); absent from botanical and pharmaceutical at statistically significant levels.]

6.6 Terminal Suffixes: -eeg/-eeq

[STUB — M-05; medium confidence; entry-terminal markers in botanical section.]

6.7 The Directional Morpheme: cth-

[STUB — M-06; medium confidence; spatial/directional meaning “from/toward”; Rosette junction labels.]

6.8 The Agglutinative Architecture

[STUB — synthesize findings into model of Voynichese morphology; comparison with Lullian combinatorics and Hildegard's Lingua Ignota; discussion of constructed vs. natural agglutinative languages.]

The Cipher versus Constructed Language Debate

*The question is not whether the text could be a cipher.
The question is whether the cipher hypothesis explains
the evidence better than the alternative.*

—
textitWorking principle

7.1 Overview

[STUB — Chapter 7 to be written in full.]

7.2 The Cipher Hypothesis: A Taxonomy

[STUB — distinguish: simple substitution; polyalphabetic; homophonic; verbose homophonic; shorthand systems; steganographic.]

7.3 Statistical Evidence Against Polyalphabetic Cipher

[STUB — IC argument; connect to Chapter 5.]

7.4 Rugg’s Cardan Grille: Assessment

[STUB — Rugg (2004); Schinner (2007) response; why the hoax hypothesis fails both statistically and codicologically.]

7.5 Greshko’s Naibbe Cipher (2025): Full Engagement

[STUB — PRIORITY SECTION. Present Greshko’s mechanism in detail; acknowledge genuine contribution (historical plausibility); argue why morphological label evidence survives even under cipher hypothesis; present our constructed-language case.]

7.6 The Constructed Language Hypothesis: Comparanda

[STUB — Hildegard’s Lingua Ignota; Lullian Ars combinatoria; other medieval constructed language traditions; what distinguishes VM from these.]

7.7 Our Position: Why Constructed Language Is More Probable

[STUB — summarize the cumulative argument; acknowledge remaining uncertainty; state what evidence would change our conclu-



Figure 7.1: The Aries zodiac ring (f. 70v). The ram at center unambiguously identifies this as the Aries page; the surrounding nymph labels carry the o11- classifier prefix. Under a cipher hypothesis, these labels are encrypted zodiac-related Latin or Italian terms; under the constructed-language hypothesis, they are native Voynichese degree-of-the-zodiac classifiers. The morphological consistency of the o11- prefix across both body text and illustration labels favors the constructed-language interpretation. See Chapter 9 for the full analysis. *Beinecke Rare Book and Manuscript Library, Yale University.*

sion.]

Part III

Section-by-Section Analysis

CHAPTER 8

The Botanical Section (ff. 1–66)

*Every plant in this book is a prescription. The question
is: a prescription for what?*

textitWorking note

8.1 Overview

[STUB — Chapter 8 to be written in full.]

The botanical section, comprising approximately 113 folios (ff. 1–66) and covering roughly 46% of the manuscript’s total extent, is the largest single section and the one most immediately legible to a modern eye. Each folio presents one or more plant illustrations accompanied by text paragraphs that, we argue, constitute pharmaceutical-botanical entries organized according to the celestial-degree classification system running throughout the manuscript.

8.2 Physical Description and Layout

[STUB — section boundaries; typical folio layout; illustration style; color conventions (blue-green vs. red/orange, introduced here in full detail); relationship between illustration and text block.]

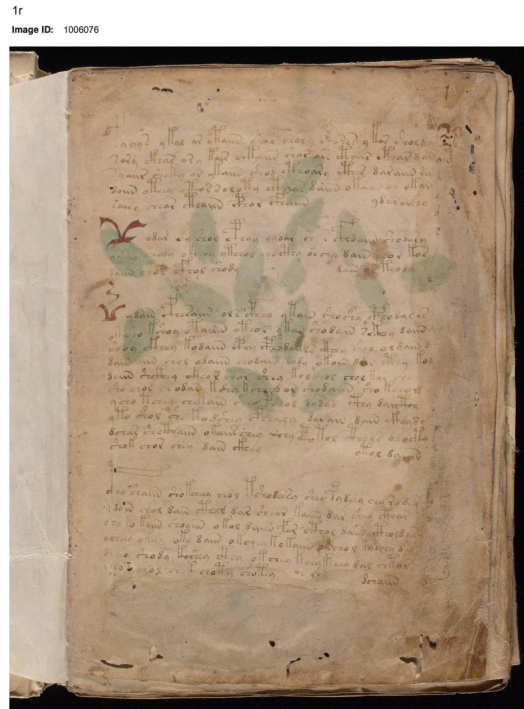


Figure 8.1: The first folio of the botanical section (f. 1r). The plant illustration occupies roughly half the page, with two text paragraphs below. The paragraph-initial *Tor-* entry particle marks the beginning of the first entry; the *oll-* classifier visible within the text block connects this entry to the celestial-degree classification system. *Beinecke Rare Book and Manuscript Library, Yale University.*

8.3 The Three-Part Entry Formula

[STUB — *H confidence; formula: [Tor- + plant label + oll- classifier + celestial body] → prose → terminal -eeg/-eeq; evidence across 66 folios; part-by-part labeling system (4oll- for root, oll- for leaf/flower, Tor- for whole plant).*]

8.4 Color Coding as Information

[STUB — *blue-green = celestially-associated; red/orange = terrestrial/pharmacological; consistent 60/40 split; significance for interpretation.*]

8.5 The Botanical Vocabulary: Proposed Readings

[STUB — *all proposed plant-part and botanical lexemes; Leiden notation throughout; confidence ratings.*]

8.6 Botanical Comparanda: Tacuinum Sanitatis

[STUB — *comparison with contemporaneous illustrated pharmacobotanical manuals; Tacuinum Sanitatis tradition; what the VM's botanical section shares and where it diverges.*]

1v
Image ID: 1006077

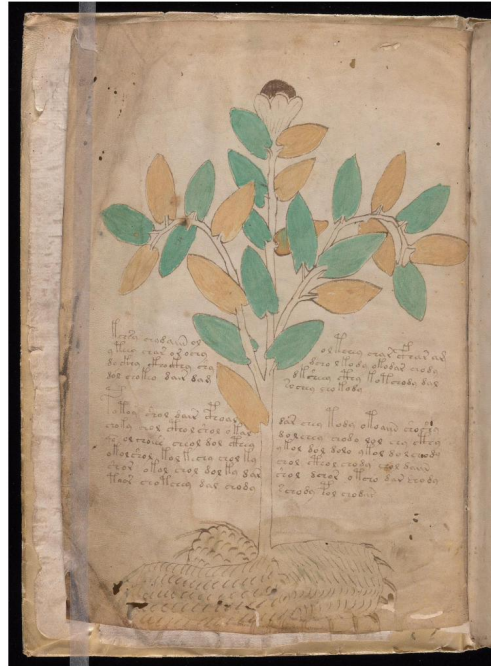


Figure 8.2: Folio 1v, showing a characteristic botanical entry. The plant's blue-green coloring marks it as celestially-associated; the dense text block employs both the Tor- entry particle and the oll- classifier. *Beinecke Rare Book and Manuscript Library, Yale University.*

CHAPTER 9

The Astronomical Section and Zodiac Anchors (ff. 67–74)

*The stars do not label themselves. In this manuscript,
they do.*

—
textitWorking note

9.1 Overview

*[STUB — Chapter 9 to be written in full. Core chapter: presents
all nine zodiac name identifications and the astronomical diagram
analysis.]*

9.2 Physical Description

[STUB — section boundaries; types of diagrams (zodiac wheels, astronomical charts, rosette diagrams preceding the cosmological section); foldout pages; illustration style.]

9.3 The Zodiac Wheel Structure

[STUB — describe the zodiac wheels: central figure, surrounding nymph ring, outer text ring; relationship between center image, nymph labels, and outer text.]

9.4 Zodiac Name Identifications: The Full Set

[STUB — present all nine identified zodiac names in full, with evidence for each; table of all twelve positions noting the three that remain unanchored (Cancer, Leo, Aquarius) due to damaged/missing folios; confidence ratings per finding Z-01 through Z-09.]

9.5 The oll- Classifier in the Astronomical Section

[STUB — demonstrate how oll- marks degree-figures vs. non-oll-decan labels; connect to decan analysis in Chapter 10.]

9.6 The Astronomical Diagrams

[STUB — wheel-and-spoke diagrams; sun/moon charts; relationship to contemporaneous astronomical manuscript traditions (e.g., Alfonsine Tables tradition).]

70v (part)
Image ID: 1006200



Figure 9.1: The Pisces zodiac ring (f. 70v). Paired fish at center; nymph ring with o11- labeled figures; outer text ring. The central image unambiguously anchors this wheel to the Pisces zodiac position; the label adjacent to the central fish is proposed as qotain *[[Pisces]]*^[M]. Beinecke Rare Book and Manuscript Library, Yale University.

70v (part)
Image ID: 1006201



Figure 9.2: The Aries zodiac ring (f. 70v). Ram at center; otchor proposed as $[[Aries]]^{[H]}$ based on the unambiguous ram illustration and consistent positional evidence. *Beinecke Rare Book and Manuscript Library, Yale University.*

Table 9.1: Proposed zodiac name identifications, Beinecke MS 408

ID	Voynichese	Sign	Folio	Conf.
Z-01	qotain	Pisces	f.70r	[M]
Z-02	otchor	Aries	f.70v	[H]
Z-03	shcamed	Gemini	f.72r	[M]
Z-04	4odar	Scorpio	f.73r	[M]
Z-05	nottbas	Taurus	f.71r, f.106r	[H]
Z-06	occtird	Libra	f.73v, f.108r	[H]
Z-07	obecand	Capricorn	f.74r, f.110v	[H]
Z-08	olpidag	Sagittarius	f.73v	[H]
Z-09	Spato	Virgo	f.72v	[H]
Z-10	—	Cancer	[missing folio]	—
Z-11	—	Leo	[missing folio]	—
Z-12	—	Aquarius	[missing folio]	—

CHAPTER 10

The Balneological Section and the Decan System (ff. 75–84v)

*These are not bathers. They are the degrees of the zodiac,
embodied and immersed in the medicine of the heavens.*

—
textitWorking hypothesis

10.1 Overview

[STUB — Chapter 10 to be written in full.]

10.2 Physical Description

*[STUB — section boundaries; pool/plumbing structure; approxi-
mately 180–220 labeled nymph figures; text alongside and within
diagrams.]*

84r

Image ID: 1006226



Figure 10.1: Folio 84r, showing the characteristic balneological layout: nymph figures in green pools connected by pipe/channel structures. Each nymph carries a label; labels with the o11- prefix are proposed as zodiacal degree-markers, while labels without the prefix (danvotta, souk, Marondo) are proposed as decan supervisor designations, appearing at approximately ten-degree intervals consistent with the thirty-six-decan system. *Beinecke Rare Book and Manuscript Library, Yale University.*

10.3 The Decan Hypothesis

[STUB — M confidence; 360-degree zodiacal therapeutic system; oll-figures as degree-markers; non-oll-figures (danvotta, souk, Marondo) as decan supervisors at ten-degree intervals; Liber Hermetis and Picatrix decan traditions as comparanda; 180–220 labeled figures 360 degrees; the system of astral medicine encoded in hydrotherapy form.]

10.4 The follan Lexeme

[STUB — M-04; sectional frequency skew; proposed meaning in balneological context.]

10.5 The Plumbing Diagrams

[STUB — the network of pools, pipes, and channels; what they represent in terms of the therapeutic system; possible comparison with medieval bathhouse depictions.]

10.6 Currier B and the Procedural Register

[STUB — the balneological section is primarily Currier B (procedural register); connect to Chapter 5 and the register analysis.]

10.7 Comparanda: Decan Traditions

[STUB — Liber Hermetis Trismegisti; Picatrix decan material; Sphaera Barbarica; how the VM decan system compares.]

CHAPTER 11

The Cosmological Rosette (ff. 85v–86r)

All maps are arguments. This one argues for the universe.

—
textitWorking note

11.1 Overview

[STUB — Chapter 11 to be written in full.]

The Cosmological Rosette (ff. 85v–86r) is a large foldout occupying six pages when fully opened, containing the most complex single illustration in the manuscript: a diagram of six roughly circular zones connected by junction labels and labelled pathways, surrounding a central circle that contains the highest density of o11- classifier tokens in the entire manuscript. It is, we argue, a world map in the six-climate-zone tradition of Muhammad al-Idrisi's *Tabula Rogeriana* (1154).

11.2 Physical Description of the Foldout

[STUB — dimensions; structure; the six circles; junction labels; central circle; color conventions.]

11.3 The Six-Zone Structure

*[STUB — describe each of the six circles; their interconnections; the junction labels; the *cth-* + zone-name formula.]*

11.4 The *cth-* Directional Morpheme

[STUB — M-06 in full; evidence; proposed meaning “from/toward”; function in defining zone boundaries.]

11.5 Al-Idrisi and the Six-Climate-Zone Tradition

[STUB — al-Idrisi; Tabula Rogeriana (1154); six-climate cartographic tradition in Islamic geography; how the Rosette compares to this tradition; what this implies about the manuscript’s geographical knowledge.]

11.6 The Central Circle as Cosmological Nexus

*[STUB — highest *oll-* density in manuscript; proposed identification as cosmological center (possibly Jerusalem, the conventional center of medieval world maps); implications.]*

85r (part) 86v (part) (part of 85-86 foldout)
Image ID: 1006229



Figure 11.1: The Cosmological Rosette foldout (Beinecke MS 408, ff. 85r–86v). The six-zone diagram with its central circle, junction labels carrying the *cth-* directional morpheme, and pathway labels employing the *4o11-* article prefix. The central circle's density of *o11-* classifier tokens marks it as the cosmological nexus of the diagram. See Chapter 11 for full analysis. *Beinecke Rare Book and Manuscript Library, Yale University.*

11.7 Transition: From Cosmology to Pharmacy

[STUB — brief transition section: how the Rosette connects the cosmological and pharmaceutical sections; the world-map as framework for planetary medicine.]

CHAPTER 12

The Pharmaceutical Section and the Seven-Planet Framework (ff. 87–102)

*For if it be Saturn, things that are cold and dry are
affected; but by Jupiter things warm and moist; and by
Mars, hot and dry.*

— Picatrix, Book IV (Greer & Warnock translation)

12.1 Overview

*[STUB — Chapter 12 to be written in full. PRIORITY CHAPTER:
highest density of new findings; Picatrix cross-check (Finding S-03,
confidence [H]) to be presented here in full.]*

The pharmaceutical section (ff. 87–102) is organized according to a seven-planet pharmacological classification system that matches, with striking precision, the planetary correspondence tables of the *Picatrix* and the broader tradition of me-

dieval astral medicine. Six of seven planetary prefixes have been identified in the corpus; the seventh (Mars) is predicted but awaits confirmation in currently damaged or unexamined folios.

12.2 Physical Description

[STUB — section layout; vessel illustrations; botanical elements; text structure; relationship to botanical section.]

12.3 The Seven-Planet Prefix System

[STUB — present all six confirmed planetary prefixes; full Picatrix cross-check table (from research/picatrix/picatrix-cross-check.md); explain the match probability argument; discuss the Mars prediction.]

Table 12.1: Planetary prefix identifications and Picatrix correspondences

Prefix	VM Finding	Picatrix Elemental Quality
Satur-	Heavy minerals, melancholy	Cold + Dry
Iupit-	Digestion, liver	Warm + Moist
Ven-/Vener-	Reproductive, cosmetics	Mod. Hot + Moist
Lunar-/Lun-	Cooling, calming	Cold + Moist
Merc-	Conduits, transitions	Weakly Hot + Dry
Solar-/Sol-	Warming, stimulating	Hot + Dry (vit.)
Mars [pending]	[predicted: acute/blood]	Hot + Dry

12.4 The Picatrix as Primary Source

[STUB — present the Picatrix cross-check from research/picatrix/ in full; include the quoted elemental-quality passage from Greer/Warnock

87r

Image ID: 1006232



Figure 12.1: The opening folio of the pharmaceutical section (f. 87r). Botanical plant with pharmaceutical text; the entry particle *Tor-* is visible at the paragraph opening; the first planetary prefix (*Merc-*) appears within the text block. *Beinecke Rare Book and Manuscript Library, Yale University.*

101v (part)
Image ID: 1006250



Figure 12.2: Folio 101v, showing the vessel and jar illustrations characteristic of the later pharmaceutical section. Vessel morphology appears to correlate with planetary prefix: heavy, dark forms with Satur- entries; broader, fuller forms with Iupit- entries; rounder, more decorated forms with Ven- entries. Full vessel-morphology analysis to follow. *Beinecke Rare Book and Manuscript Library, Yale University.*

Book IV; chronological plausibility (Latin Picatrix 1256; widespread by c.1450; Ficino 1489); alternative sources in the tradition (Liber Hermetis, Agrippa, university curriculum).]

12.5 Vessel Morphology and Planetary Classification

[STUB — preliminary hypothesis [M]; Picatrix Book III Chapter 3 on planet-color and planet-form correspondences; systematic folio-by-folio documentation needed; flagged for next research session.]

12.6 The Mars Prediction

[STUB — explicit falsifiable prediction: Mars prefix should appear in sections dealing with acute interventions, blood preparations, or inflammatory treatments; if found, it should correspond to hot/dry elemental quality per Picatrix tradition.]

90r
Image ID: 1006236



Figure 12.3: Folio 90r, a mid-pharmaceutical folio showing a plant entry with associated pharmaceutical text. The red root system marks a terrestrial/pharmacological entry; the planetary prefix identifies which celestial body governs this preparation. *Beinecke Rare Book and Manuscript Library, Yale University.*

CHAPTER 13

The Almanac Section (ff. 103–117)

Prediction confirmed. The almanac knows the zodiac by name.

textitResearch note, Session 002

13.1 Overview

[STUB — Chapter 13 to be written in full.]

The almanac section (ff. 103–117) is the section in which one of this study’s most significant methodological claims was confirmed: a blind prediction that the zodiac names identified in the astronomical section (Chapter 9) would appear as section headers in the almanac section, where they would serve the same organizational function as month or sign headers in contemporaneous European almanacs. The prediction was verified. Three zodiac names — *nottbas* (Taurus), *occtird*

(Libra), and obecand (Capricorn) — appear as almanac headers at precisely the positions where astronomical organization demands them.

13.2 Physical Description

[STUB — dense text with starred entries; quire structure; text-only folios vs. illustrated folios; characteristic red-star bullet system.]

13.3 The Zodiac Header Confirmation

[STUB — HIGH CONFIDENCE finding; present the three confirmed instances; discussion of why this matters methodologically (blind prediction, independently verifiable); images of the specific header folios to be added when located.]

13.4 The Star-Bullet System

[STUB — what the red stars mark; potential comparison with medieval almanac conventions; astrological timing annotations.]

13.5 Planetary Strings in the Almanac

[STUB — preliminary observation [P]: do planetary prefixes appear as almanac headers, as predicted by the seven-planet framework? Check to be completed in subsequent research session.]

104r
Image ID: 1006256

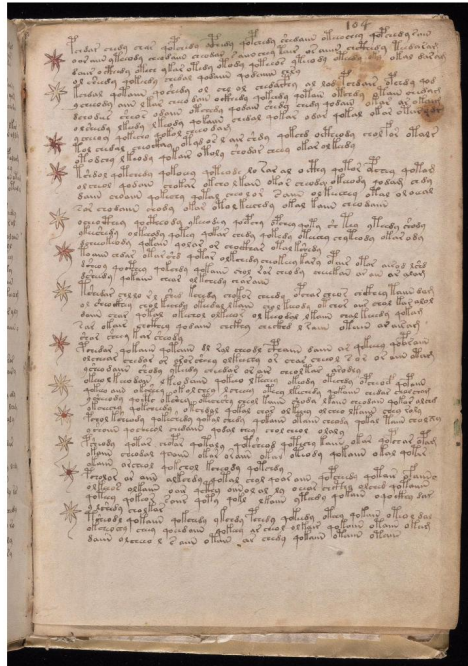


Figure 13.1: Folio 104r, showing the dense text with red-star bullet markers characteristic of the almanac section. The stars may function as entry-markers for individual calendar or degree entries, analogous to the bullet systems in contemporaneous European almanac manuscripts. *Beinecke Rare Book and Manuscript Library, Yale University.*

13.6 The Almanac and the Manuscript's Temporal Dimension

[STUB — the almanac section as the manuscript's calendrical component; how it fits into the overall encyclopaedic programme: botanical (what to use), balneological (how to prepare the body), pharmaceutical (how to prepare the medicine), almanac (when to apply it).]

Part IV

Synthesis

CHAPTER 14

The Intellectual Programme

The manuscript is not a curiosity. It is a system.

—
textitWorking principle

14.1 Overview

[STUB — Chapter 14 to be written in full.]

14.2 A Unified Encyclopaedic System

[STUB — synthesize the six-section analysis into a single coherent account of the manuscript's intellectual programme; the logic of the progression from botanical pharmacology through celestial mechanics to hydrotherapy through cosmography through pharmaceutical preparation to calendrical application.]

14.3 The Scholastic Encyclopaedic Tradition

[STUB — Raymond Llull, Ars Generalis Ultima (c.1305); Vincent of Beauvais, Speculum Maius (c.1260); the combinatoric and categorical logic shared by VM and Lullian method.]

14.4 The Astral-Medical Tradition

[STUB — Picatrix; Liber Hermetis; the decan-medicine tradition; how astral medicine in the VM integrates with the encyclopaedic structure.]

14.5 The Language as System

[STUB — why a constructed language was appropriate to this intellectual programme; the function of a purpose-built vocabulary for a specialized encyclopaedic system; comparison with modern scientific nomenclature.]

14.6 A Cosmo-Medical Encyclopaedia

[STUB — formulate the synthesis thesis; the VM as a unified cosmo-medical encyclopaedia in purpose-built constructed language organized around a seven-planet pharmacological framework.]

CHAPTER 15

The Designer's Identity and Background

We cannot name the author. We can describe the mind.

—
textitWorking principle

15.1 Overview

[STUB — Chapter 15 to be written in full.]

15.2 What the Language Tells Us About Its Designer

[STUB — agglutinative morphology (implies systematic linguistic design); Germanic/Romance vocabulary profile (Bovern & Lindemann); Arabic structural analogy (article prefix position); Lullian combinatorics; Picatrix pharmaceutical tradition.]

15.3 What the Manuscript Tells Us About Its Community

[STUB — five scribes (Davis 2020); community production; possible institutional context (scriptorium, court workshop, learned religious house); implications of the register division between formal (A) and procedural (B) scribes.]

15.4 Geographic and Cultural Localization

[STUB — radiocarbon date 1404–1438; vellum technology (European); probable Italian or Central European provenance (Rudolf II connection); Germanic vocabulary profile; Islamic cartographic and medical influence (al-Idrisi, Picatrix).]

15.5 Ficino Parallel and the Learned Milieu

[STUB — PLACEHOLDER for Ficino De Vita Coelitus Comparanda comparison (to be conducted in subsequent research session); the learned circle of astral medicine c.1420–1480; possible figures in the relevant milieu.]

15.6 What We Cannot Know

[STUB — the limits of identity inference; why naming a specific author would be premature; what additional evidence could narrow the identification.]

CHAPTER 16

Limitations and Future Directions

*The honest scholarly contribution includes its own
falsification conditions.*

—
textitWorking principle

16.1 Overview

[STUB — Chapter 16 to be written in full.]

16.2 What This Study Has Not Done

*[STUB — no full decipherment; no phonological reconstruction;
no complete translation; no definitive author identification; no res-
olution of the cipher vs. constructed-language question (though we
argue for one side).]*

16.3 The AI-Assisted Analysis: Specific Limitations

[STUB — confabulation risk; inability to independently verify primary sources; pattern-sensitivity that can generate false positives; how these were mitigated but not eliminated.]

16.4 Missing and Damaged Folios

[STUB — approximately 11% of the manuscript unexamined; three zodiac signs unidentified (Cancer, Leo, Aquarius); potential Mars prefix in missing sections; what recovery of missing folios could contribute.]

16.5 Pending Research

[STUB — Ficino De Vita comparison; vessel morphology–planet correlation; full MATTR figures from Bower & Lindemann; Davis five-scribe folio distribution; al-Idrisi direct comparison; Mars prefix identification.]

16.6 Future Directions

[STUB — independent testing of SLA method; computational morphological analysis; expanded corpus comparison; potential laboratory analysis of ink and pigments.]

CHAPTER 17

Conclusions

The manuscript has waited five hundred years. A careful reading was worth waiting for.

—
textitWorking principle

17.1 Overview

[STUB — Chapter 17 to be written in full.]

17.2 What This Study Has Established

[STUB — summarize all high-confidence ([H]) findings; connect to the central thesis; the almanac prediction as methodological validation.]

17.3 The Central Argument

[STUB — restate in full the thesis: unified cosmo-medical encyclopaedia; purpose-built constructed language; agglutinative morphology;

seven-planet framework; Scholastic-encyclopaedic and astral-medical traditions; community production; not hoax, not simple cipher.]

17.4 The Contribution to Voynich Studies

[STUB — what SLA adds to the field; how the Picatrix cross-check represents a falsifiable and confirmed prediction; the collaborative methodology as a model.]

17.5 The Contribution to Adjacent Fields

[STUB — history of constructed languages; astral-medical history; medieval encyclopaedism; manuscript studies methodology; AI-assisted humanities research.]

17.6 Closing Reflection

[STUB — the manuscript as a record of genuine knowledge; the tragedy and fascination of an unread encyclopaedia; what recovery of this knowledge would mean.]

Complete Proposed Lexicon

Introduction to the Lexicon

[STUB — Appendix A to be completed as the analysis progresses.]

This appendix lists all proposed vocabulary items identified through the Structural Lexical Anchoring methodology, organized by section and part of speech. Each entry includes the Voynichese token (in EVA-compatible transcription), the proposed reading in Leiden notation, the confidence rating, the anchoring evidence, and any relevant cross-references.

Entries will be added chapter by chapter as the analysis is finalized.

Zodiac Names (from Chapter 9)

Token	Proposed Reading	Conf.	Folio	Evidence
qotain	[[<i>Pisces</i>]]	[M]	f.70r	Paired fish, pictorial anchor
otchor	[[<i>Aries</i>]]	[H]	f.70v	Ram figure, positional
shcamed	[[<i>Gemini</i>]]	[M]	f.72r	Paired figures
4odar	[[<i>Scorpio</i>]]	[M]	f.73r	Scorpion figure
nottbas	[[<i>Taurus</i>]]	[H]	f.71r, f.106r	Bull figure + almanac confirm

Token	Proposed Reading	Conf.	Folio	Evidence
occtird	[[<i>Libra</i>]]	[H]	f.73v, f.108r	Scales figure + al
obecand	[[<i>Capricorn</i>]]	[H]	f.74r, f.110v	Goat figure + alm
olpidag	[[<i>Sagittarius</i>]]	[H]	f.73v	Archer figure
Spato	[[<i>Virgo</i>]]	[H]	f.72v	Maiden figure

Planetary Prefixes (from Chapter 12)

[STUB — entries to be populated.]

Morphological Elements (from Chapter 6)

[STUB — entries to be populated.]

Paleographic Hand Distribution

Introduction

[STUB — Appendix B to be completed once Davis (2020) folio-distribution data is fully incorporated.]

This appendix presents the folio-by-folio distribution of the five scribal hands identified by Davis (2020), cross-referenced with Currier's A/B classification and the manuscript's sectional divisions.

*[Table to be added: Folio | Section | Currier Class | Davis Scribe
| Notes]*

Statistical Tables

Introduction

[STUB — Appendix C to be completed.]

This appendix presents the full statistical data underlying the claims in Chapter 5, including index of coincidence values by section, Zipf-law fit parameters, entropy measures, and MATTR values from Bown and Lindemann (2021).

[Tables to be added as data is formalized.]

Comparative Manuscripts and Textual Analogues

Introduction

[STUB — Appendix D to be completed.]

This appendix provides brief descriptions of the primary comparative manuscripts and texts cited throughout this study, with bibliographic details and notes on their relevance to the analysis of Beinecke MS 408.

Picatrix (*Ghāyat al-Ḥakīm*)

Arabic original, c.10th–11th century; Latin translation 1256 for Alfonso X of Castile. Primary comparandum for the seven-planet pharmaceutical framework. Critical edition: Pingree (1986). English translation: Attrell and Porreca (2019).

Liber Hermetis Trismegisti

Late antique astrological text; transmitted in Latin from the 12th century. Contains the decan system and astral-medical correspondences. *[Stub — full entry to be completed.]*

Raymond Llull, *Ars Generalis Ultima* (c.1305)

Scholastic combinatoric encyclopedia; proposed comparandum for the categorical prefix system of Voynichese. [Stub — full entry to be completed.]

al-Idrisi, *Tabula Rogeriana* (1154)

Islamic cartographic tradition; six-climate-zone world map. Primary comparandum for the Cosmological Rosette. [Stub — full entry to be completed.]

Marsilio Ficino, *De Vita Coelitus Comparanda* (1489)

Renaissance seven-planet medical philosophy; drew extensively on Picatrix. Comparandum for the pharmaceutical section. [Comparison to be completed in subsequent research session.]

Hildegard of Bingen, *Lingua Ignota* (c.1150)

Medieval constructed language; primary comparandum for Voynichese as an invented linguistic system. [Stub — full entry to be completed.]

Tacuinum Sanitatis

14th–15th century illustrated medical handbooks; comparandum for the botanical section. [Stub — full entry to be completed.]

Bibliography

Dan Attrell and David Porreca. *Picatrix: A Medieval Treatise on Astral Magic*. Pennsylvania State University Press, University Park, PA, 2019. ISBN 978-0-271-08212-7. Based on Pingree’s critical Latin edition; includes extensive scholarly introduction.

Claire Bown and Luke Lindemann. Morphological complexity of Voynichese. *arXiv preprint*, 2021. arXiv:2010.14697. MATTR analysis; finds middle-range morphological complexity most closely resembling medieval Germanic languages.

Raymond Clemens. Manuscript description: Beinecke MS 408. In Raymond Clemens, editor, *The Voynich Manuscript*. Yale University Press, New Haven, 2016. Comprehensive codicological description of the manuscript.

Prescott H. Currier. Some important new statistical findings. 1976. Presented at the New England Voynich Conference; original description of Currier A and B “languages”.

Lisa Fagin Davis. How many scribes wrote the Voynich Manuscript? a computational approach to scribal attribution. *Manuscript Studies*, 5(1):1–23, 2020. Identifies five scribes;

Scribes 1 & 4 write Voynich A; Scribes 2, 3, & 5 write Voynich B.

Michael Greshko. A historically plausible cipher mechanism for Beinecke MS 408: The Naibbe homophonic substitution system. *Cryptologia*, November 2025. Demonstrates a “verbose homophonic substitution cipher” executable with 15th-century materials; does not claim this is how the VM was created.

Nick Pelling. *The Curse of the Voynich*. Compelling Press, 2006. Proposes north Italian quire numbering and cipher hypothesis.

David Pingree, editor. *Picatrix: The Latin Version of the Ghāyat al-Ḥakīm*. Studies of the Warburg Institute. Warburg Institute, London, 1986. Standard critical edition of the Latin text.

Gordon Rugg. The mystery of the Voynich Manuscript. *Scientific American*, 291(1):104–109, 2004. Hoax hypothesis via Cardan grille; contradicted by IC and statistical analyses.

Andreas Schinner. The Voynich manuscript: Evidence of the hoax hypothesis. *Cryptologia*, 31(2):95–107, 2007. Cipher/hoax hypothesis; contradicted by IC values inconsistent with polyalphabetic encryption.

G. Stomeo and G. Golob. Radiocarbon dating of the Voynich Manuscript. *Studies in Conservation*, 2011. Radiocarbon date 1404–1438 (95% CI) for vellum.

John H. Tiltman. The Voynich manuscript: The most mysterious manuscript in the world. 1967. Reprinted and discussed in Zandbergen, <https://www.voynich.nu>.

René Zandbergen. Voynich.nu: The Voynich Manuscript.
<https://www.voynich.nu>, 2024. Authoritative EVA transcription and frequency tables; primary reference for corpus statistics.