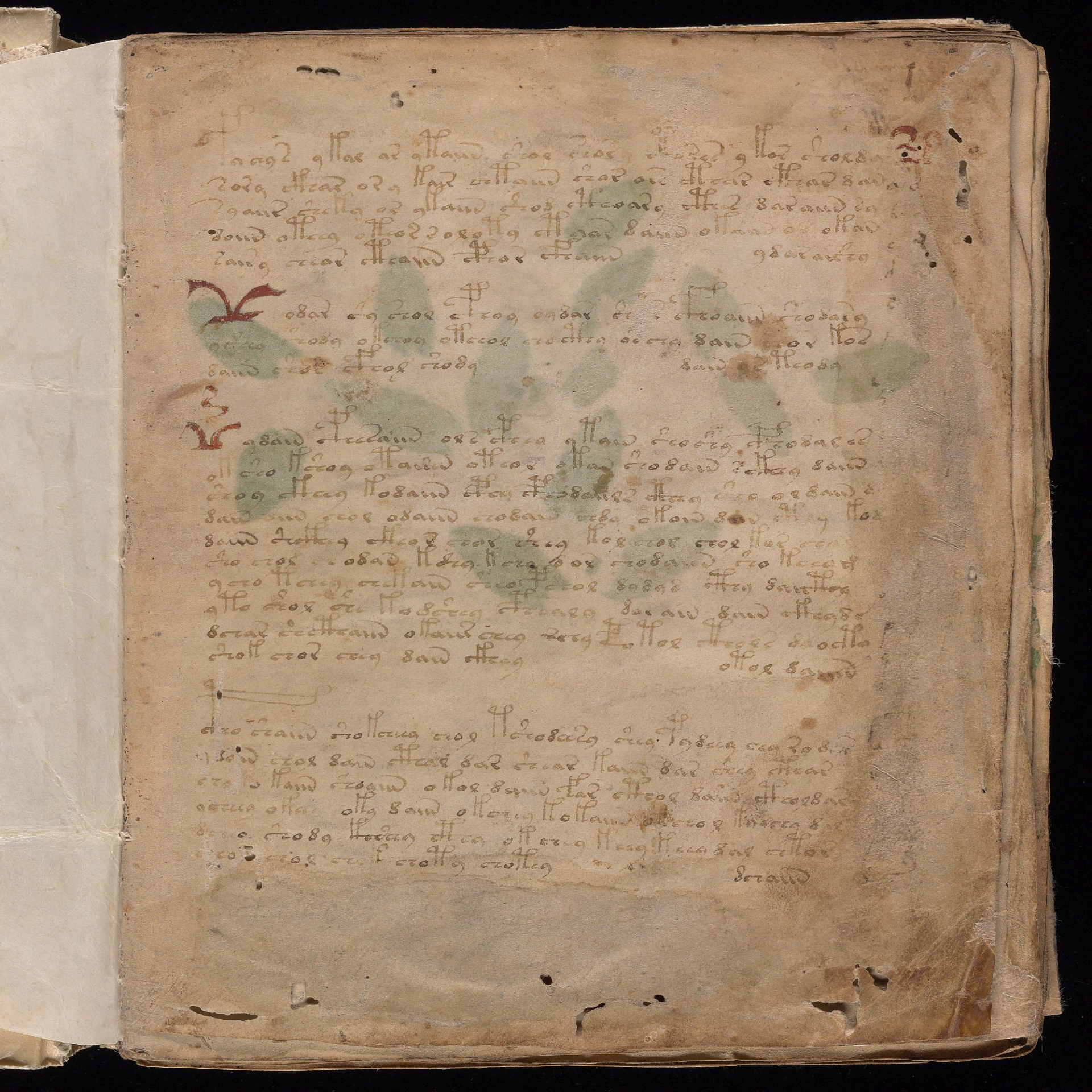
Reading Without a Key



Reading Without a Key

THE VOYNICH MANUSCRIPT, EXAMINED

What the object, the script and the numbers can tell us about a book no one has read in six hundred years

Kyle Prier

with Claude (Anthropic)

Copyright © 2026 Kyle Prier. Working draft, version 1.0, October 2026.

All images of Beinecke MS 408 are reproduced from the digital facsimile published by the Beinecke Rare Book & Manuscript Library, Yale University (`https://collections.library.yale.edu/catalog/2002046`). The manuscript is in the public domain.

All statistics in this book were computed from the Zandbergen–Landini transliteration (ZL3b, 13 May 2025, IVTFF 2.0) distributed at `https://www.voynich.nu`, with the scripts printed in Appendix A. Figures were produced with those scripts.

This book was written by a human author working with an AI system. Appendix D describes how, and what each contributed.

Set in Spectral and Inter. Typeset with XeLaTeX.

Contents

Prologue: the first page viii

I The Object 1

1 A Book Nobody Can Read 2

2 Skin, Ink and Pigment 6

3 Where It Has Been 10

4 In What Order? 14

II The Script 18

5 Writing It Down 19

6 Every Letter Knows Its Place 23

7 The Line and the Paragraph 27

8 Two Languages, Five Hands 31

9 The Numbers 36

IIIThe Pictures 41

10 Plants 42

11 Stars and the Zodiac 47

12 The Bathers 53

13 The Rosettes 57

14 Jars and Roots 60

15 Starred Paragraphs 64

IV Explanations 68

16 What It Could Be 69

17 Attempting a Reading 73

18 A Defensible Reading 78

19 Judging the Next Solution 81

V The Attack 84

20 A Framework for Testing 85

21 One Hundred Languages 88

22 Borrowed Instruments 92

23 Brute Force 97

24 Pictures as Data 103

VI The Procedure 111

25 Five Hands, One Procedure 112

26 Fragments 118

27 Lists and Alphabets 122

VII The Chancery Key 125

28 Chancery Keys 126

29 Looking for the Key 131

30 Where This Leaves the Reading 139

A Data and Methods 143

B Folio Concordance 147

C Statistical Tables 150

D How This Book Was Made 153

E Hypothesis Registry 155

F Methods Inventory 161

Bibliography 164

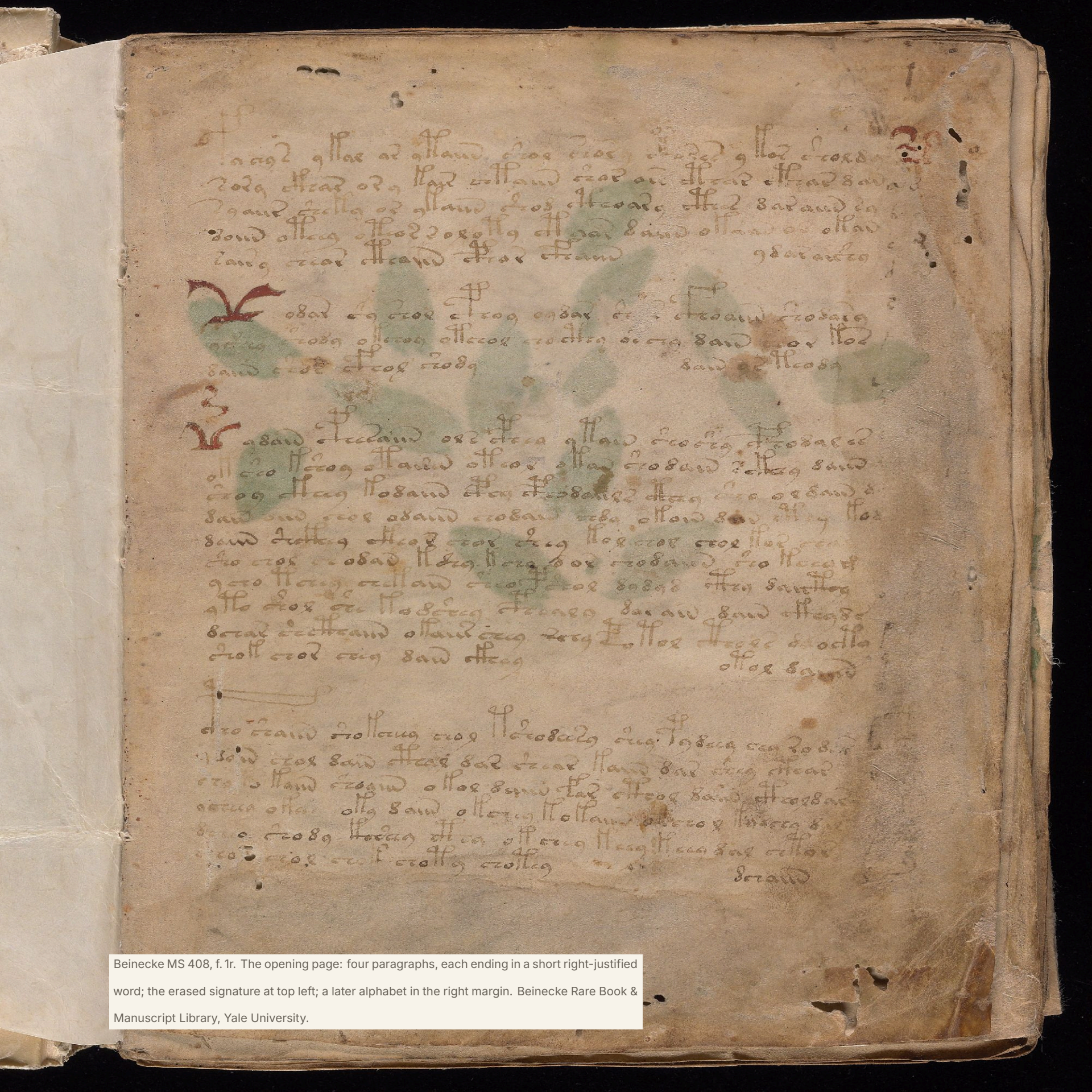
Prologue: the first page

Folio 1 recto is the page you see when you open the book, and it is not a promising start. There is no title, no decorated initial, no owner's crest. The top left corner is scuffed where Wilfrid Voynich, not long after he bought it, tried to bring up an erased signature with a chemical reagent and made it worse. Four paragraphs of text fill the page, separated by small gaps. The first begins with a tall, looping character that reaches up above the line like a gallows; the second and third begin with thick, odd marks that occur nowhere else in the book; the last begins with a gallows again. Each paragraph ends with a short word or phrase pushed to the right margin, the way a quotation might end with the name of its author. In that same margin, in a different and later hand, someone has written a column of ordinary Latin letters, *a, b, c*, as if trying a key that did not fit.

Everything about the layout says "book." The margins are ruled by eye but kept. The lines are level. The words are separated by spaces of the kind a scribe leaves without thinking. The ink is the brown of iron-gall ink that has aged on calfskin. If you have handled fifteenth-century manuscripts, your hands know what this is before your eyes do.

Then you look at a word. The first one on the page is transcribed, in the alphabet scholars use for this script, as **fachys**. Nobody knows how to say it. Nobody knows whether it is a word at all, in the sense of a unit of sound that means something, or a group of symbols standing for something else, or nothing. The second word is **ykal**; the third is **ar**; the fourth is **ataiin**, and **ataiin** is the first hint that the text has rules, because that ending, **-iin**, will turn out to close more than four thousand of the words in the book, and to appear anywhere but at the end only about a hundred times.

This book is about those rules, and about everything else that can be established about the object without being able to read it. It is written in the belief that the right response to an unsolved problem is not a solution but an inventory: what is known, how well, and from what evidence. The Voynich Manuscript has attracted more proposed solutions than any other text of its size, and nearly all of them have failed in the same way, by starting from an answer and working backwards. We start from the page.



Beinecke MS 408, f. 1r. The opening page: four paragraphs, each ending in a short right-justified word; the erased signature at top left; a later alphabet in the right margin. Beinecke Rare Book & Manuscript Library, Yale University.

Part I

The Object

1

A Book Nobody Can Read

The first thing to say about the Voynich Manuscript is how ordinary it looks. It is a small, fat book of calfskin, about the size of a modern hardback novel, in a limp vellum cover that was put on it a few centuries after it was made. Open it anywhere in the first half and you find what you would find in a hundred other fifteenth-century books: a plant, drawn in brown ink and washed with green and a muddy red, with a block of text wrapped around it. The text runs left to right. The lines are straight, the margins are tidy, the hand is fluent. It is only when you try to read a word that the ordinariness ends.

The script has never been read. Not in 1639, when a Prague alchemist named Georg Baresch wrote to the most famous scholar in Europe asking for help with it; not in 1912, when the book dealer Wilfrid Voynich bought it from a Jesuit college outside Rome and began showing it to anyone who would look; not in the century since, during which it has been attacked by the cryptanalysts of two world wars, by linguists, botanists, astronomers and historians of medicine, and by a very large number of people with a theory and an afternoon. Yale University, which has owned it since 1969, catalogues it as *Cipher manuscript* and gives its language as "Undetermined" (Beinecke Rare Book and Manuscript Library, 2020).

This book is about what can actually be known about it. That turns out to be a surprising amount, provided one is strict about the difference between what the object shows, what the text measurably does, and what one would like either of them to mean. The calfskin can be dated: it was prepared between about 1404 and 1438. The quires can be

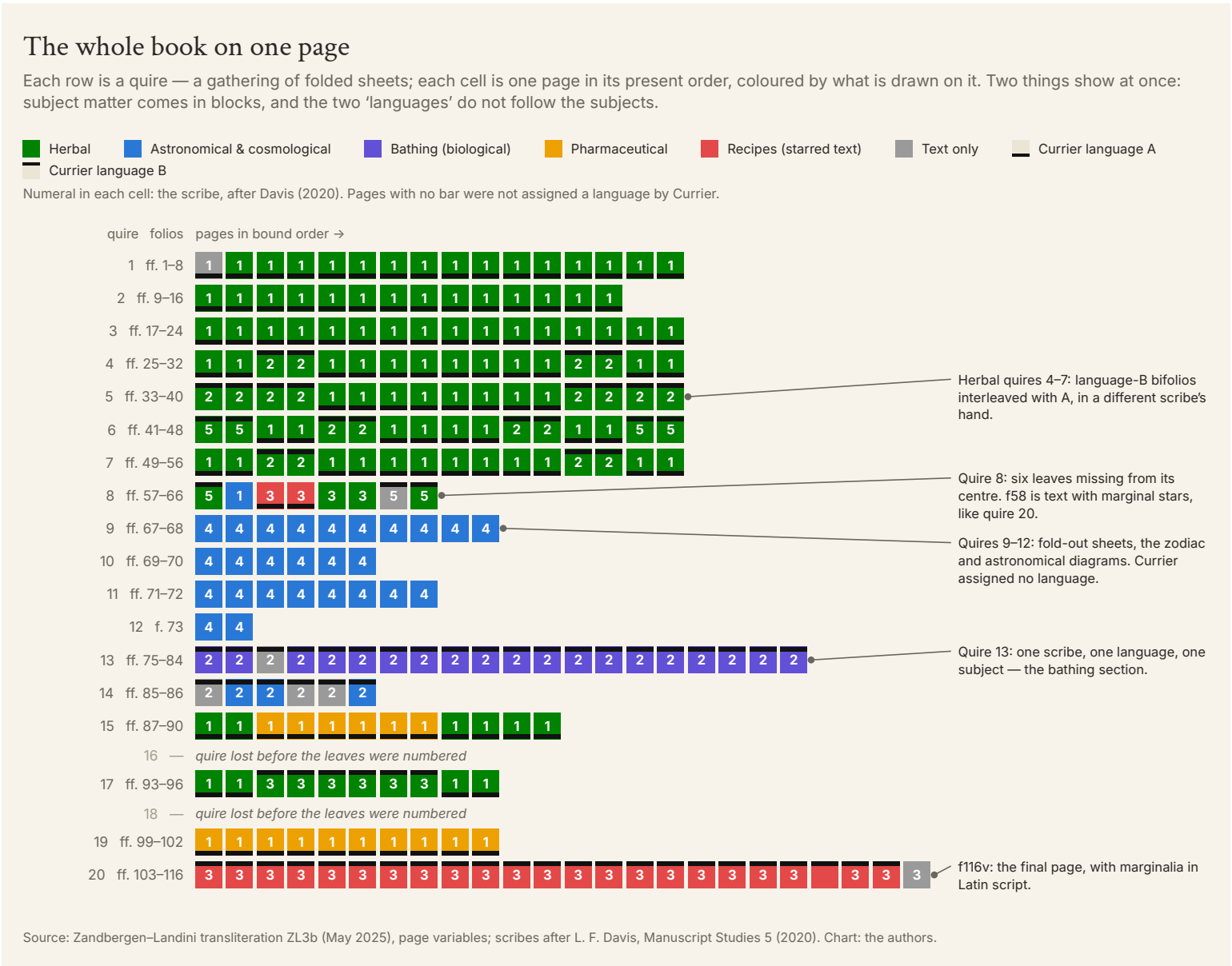


Figure 1.1 The manuscript as a grid.

counted and their original order partly recovered. The script can be transliterated, and once it is, the text can be measured in the same way any text can—how often each word recurs, how predictable each character is from the one before, how a line begins and how it ends—and those measurements place firm limits on what it could be. None of this reads the book. All of it narrows the question.

WHAT THIS BOOK DOES NOT CLAIM

It is worth being explicit, because the Voynich literature is full of books that are not. We do not offer a

decipherment. We do not identify the language, the author, the place of writing, or the purpose. Where we describe a plant as a plant or a bather as a bather we mean only that this is what the drawing shows; we do not claim to know what it is for. We use the word "word" for a group of symbols between two spaces, and "letter" or "glyph" for a symbol, without implying that either behaves like its namesake in a language.

Every number in the chapters that follow was computed from a published machine-readable transcription of the manuscript, the Zandbergen–Landini transliteration, using short programs that are printed in Appendix A and that anyone can run (Zandbergen, 1996–2026). Every statement about the physical object rests on the published conservation studies, the radiocarbon result, and the digitised folios, which Yale has placed online at full resolution and which are reproduced throughout this book (Clemens, 2016; Hodgins, 2016; Zyats et al., 2016). Where we rely on someone else's observation we say so. Where we are guessing we say that too.

WHAT IT DOES CLAIM

What we offer is a reading in the older sense of the word: a careful look at what is there. The manuscript has a structure at every level, from the stroke to the quire, and that structure is evidence. Some of it was noticed long ago. Prescott Currier, a US Navy cryptanalyst, saw in the 1970s that the text comes in two statistical dialects and that the line of writing behaves as a unit (Currier, 1976). John Tiltman, who had broken Japanese military codes, saw in the 1950s that the words are built from a small number of parts in fixed order (Tiltman, 1967). Lisa Fagin Davis, a palaeographer, established in 2020 that five people wrote it, and in 2026, with Colin Layfield, that its leaves were almost certainly bound in the wrong order by someone who could not read them (Davis, 2020; Layfield and Davis, 2026). Claire Bowern and Luke Lindemann, linguists, showed in 2021 that at the level of characters the text is more predictable than any of the 250 languages they compared it with, and at the level of words and lines looks like ordinary language (Bowern and Lindemann, 2021).

We have redone the measurable parts of that work from the current transcription, added several tests of our own, and arranged the whole into an argument. The argument runs as follows.

The object is a genuine early-fifteenth-century book, made with care by a small group of people working to a plan (Part I). The script is a real writing system in the sense that it has consistent rules, and those rules are unlike those of any known alphabet: nearly every symbol has a fixed position within the word, words are built from a few slots, the line is a unit, and the text has far less character-level variety than any European language written plainly (Part II). The pictures are of recognisable medieval genres—a herbal, a zodiac calendar, a bathing or medical sequence, a cosmological diagram, a pharmacy—drawn by people who knew the conventions but did not follow any surviving model closely (Part III). Taken together, these facts exclude several popular explanations outright and leave a small number standing, of which the strongest is that the text encodes something, by a method that compresses the

alphabet rather than expanding it, and that no one has yet reversed (Part IV).

That is not a solution. It is a map of where a solution would have to be.

HOW TO USE THIS BOOK

The chapters can be read in order or dipped into. Readers who want the pictures first should start with Part III, which can be read without Part II. Readers who want to know why every proposed decipherment so far has failed should go to chapter 18, which sets out the tests any solution has to pass, and which, we hope, will save some people a great deal of time.

Throughout, Voynichese words are given in the EVA transliteration (Landini and Zandbergen, 1998), in which the first word of the manuscript is **fachys**. EVA is a convention for writing the glyphs with Latin letters so that they can be typed and counted; the letters say nothing about pronunciation. Folio references follow Yale's foliation: f. 1r is the front of the first leaf, f. 1v its back. The large fold-out leaves have panels, numbered f. 70v1, f. 70v2 and so on, following the convention used by Zandbergen.

2

Skin, Ink and Pigment

BEFORE anyone asks what the book says, it is worth asking what it is made of, because the materials answer two questions that the text cannot: when, and whether it is what it appears to be.

THE LEAVES

The text block consists of 102 leaves of parchment, with one modern paper leaf at the back, foliated 1 to 116 with gaps (Beinecke Rare Book and Manuscript Library, 2020). A standard leaf measures about 225 by 160 millimetres, which makes a standard sheet, folded once down the middle to give two leaves, about 320 by 230. The leaves are gathered into eighteen surviving quires, numbered 1 to 20; quires 16 and 18 are missing entirely, and six further sheets are missing from the centres of quires 8 and 20 and from two other places, so that fourteen of the original 116 folios are gone (Zandbergen, 1996–2026). Most quires are the ordinary kind, four sheets nested to give eight leaves, but the book is unusual in the number of leaves that are wider than standard and fold out: some to three panels, some to four, one to six, and one sheet, the famous "rosettes" page, to roughly 45 by 45 centimetres with an extra horizontal fold. Fold-outs are not rare in medieval books, but this many in a book this small is.

The parchment is calfskin. That was established in 2014 by protein analysis at the University of York, which can tell the species of animal from a sample of collagen too small to see (Zandbergen, 1996–2026). "Vellum," strictly, means calf, so for once the common name is right.

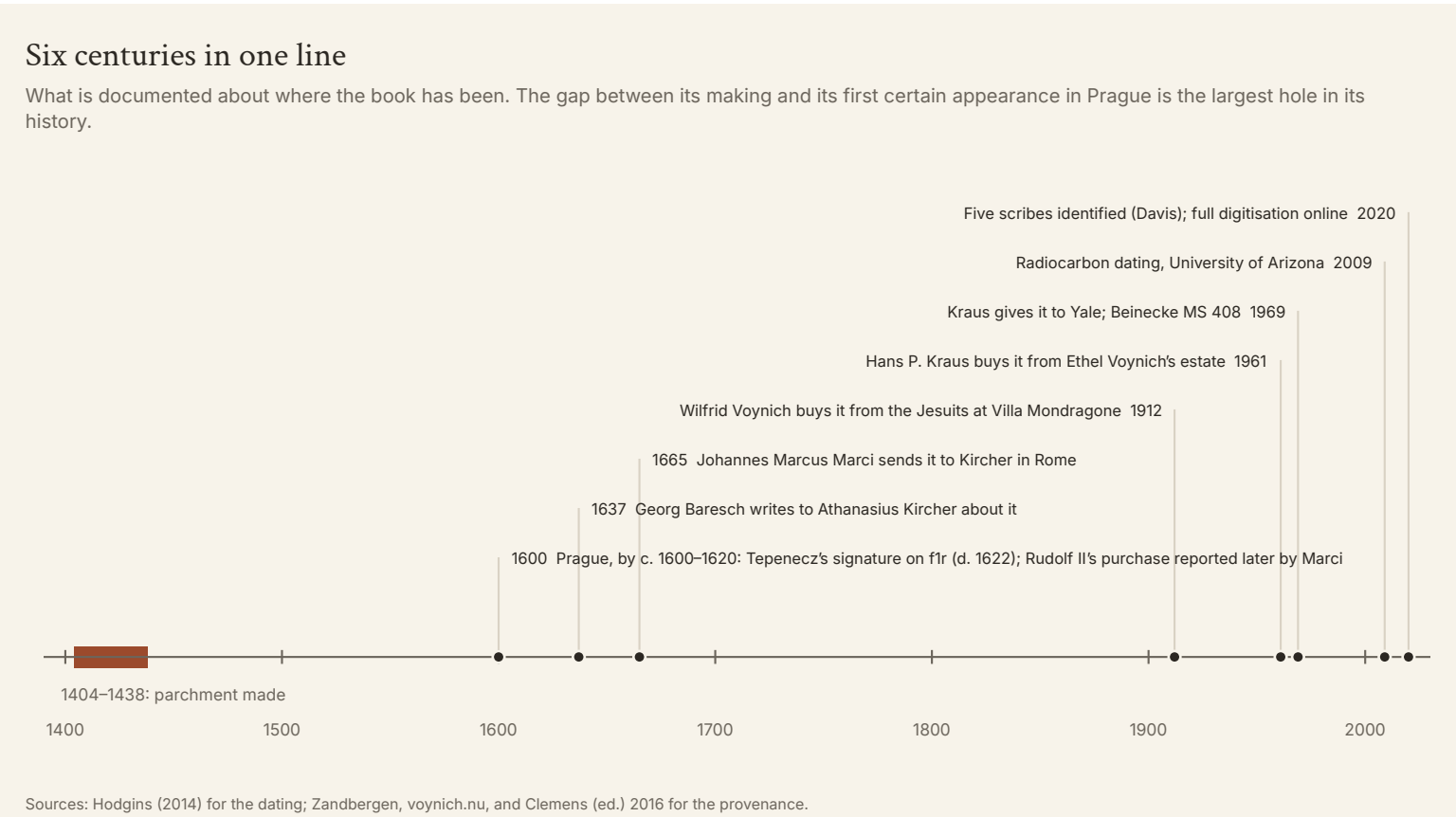


Figure 2.1 What is documented about the book's history.

THE DATE

In 2009 Greg Hodgins of the University of Arizona's accelerator mass spectrometry laboratory took four small samples from four different leaves of the manuscript and dated them by radiocarbon. The four results agree with one another, and their combined calibrated range is 1404 to 1438 at 95 per cent confidence (Hodgins, 2016). That is the date at which the calves died and the skins were made into parchment. Parchment was expensive and was not usually stockpiled for decades, so the writing is most likely from the same generation; but the date is a floor, not a ceiling. What the radiocarbon result rules out is any theory that places the book's manufacture in the sixteenth century or later, including every version of the story in which the Elizabethan magus John Dee, or Wilfrid Voynich himself, made it.

The date is consistent with everything else. The quire numbers, written in the lower corners of the last leaf of each gathering, use a form of numeral and an abbreviation for the Latin ordinal (*primus*, *2us*) that belong to the fifteenth century (Davis, 2025; Zandbergen, 1996–2026). The drawings of the zodiac, especially the crossbow-carrying Sagittarius, match conventions found in German manuscripts of roughly 1400 to 1500 (Zandbergen, 1996–2026). The

women's headdresses and the shape of the few recognisable objects in the drawings point the same way.

THE INK AND THE PAINT

The text and the outlines of the drawings are in an iron-gall ink, the ordinary writing ink of medieval Europe, made from oak galls and iron salts. Analysis of the ink and pigments was carried out for Yale by McCrone Associates in 2009 and summarised in the 2016 Yale volume on the manuscript (Zyats et al., 2016). The blue is azurite, a copper carbonate mineral; the greens are copper-based; the red-brown is an iron-oxide earth. These are the cheapest and commonest pigments of the period. There is no gold, no expensive lake, no ultramarine, nothing a wealthy patron would have paid for. A yellow wash and some brown washes are also present. The ink and the paint are consistent with the date and with each other, and nothing about them is anachronistic.

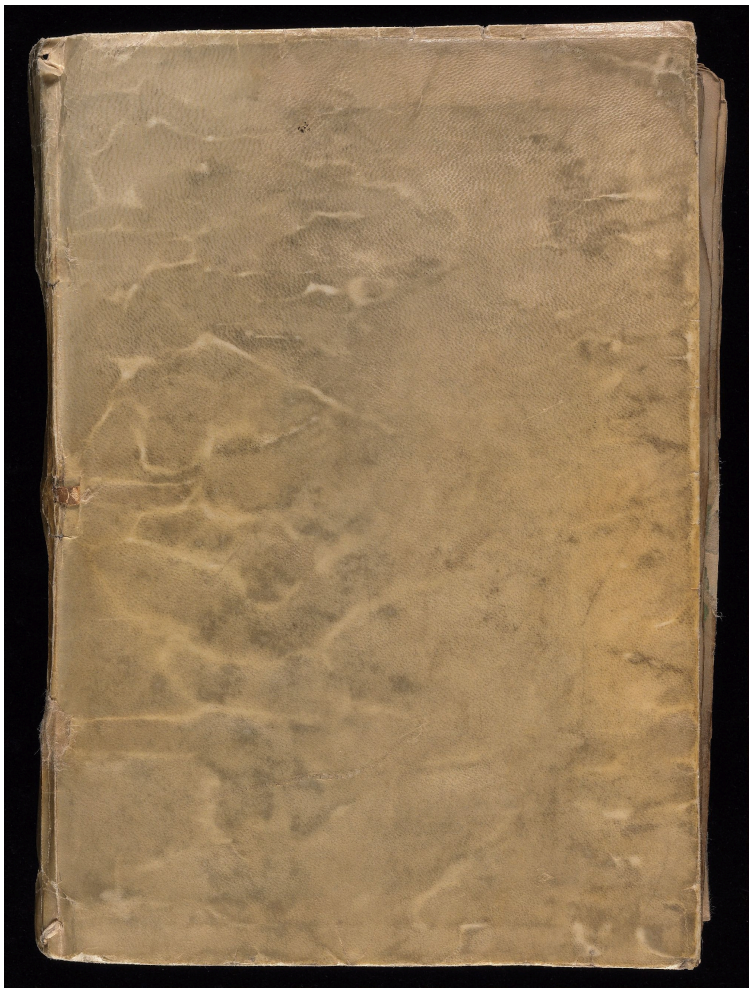
Two further observations about the paint matter for what follows. The first is that the paint was applied after the ink, in at least two campaigns, one careful and one much less so; several of the plants have been coloured by someone who did not keep inside the lines, and in a few places paint has been transferred to the facing page while still wet (Zandbergen, 1996–2026). The second follows from the first: because some of those wet-paint transfers match the page that currently faces them, the clumsy colouring was done after the leaves had reached their present order. We return to this in chapter 4, because it bears on when the book was disordered and by whom.

THE COVER AND THE SEWING

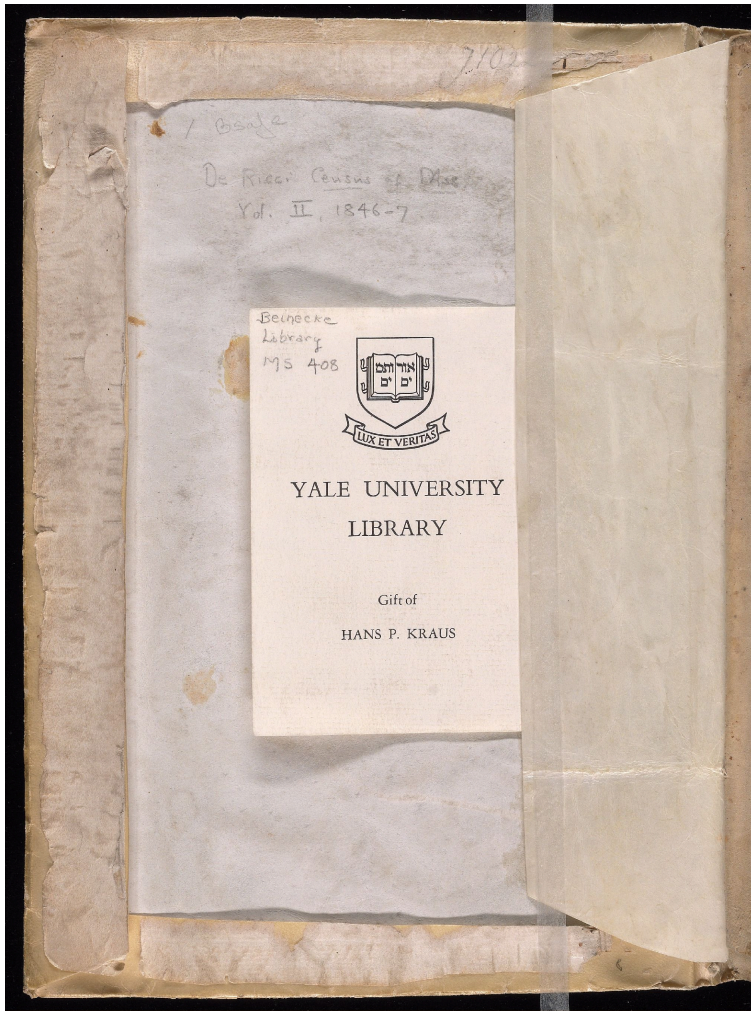
The cover is a limp vellum case of the late eighteenth or nineteenth century, probably put on in Rome by the Jesuits, who owned the book by then (Beinecke Rare Book and Manuscript Library, 2020; Davis, 2025). It is not original, and the Yale catalogue notes that it was once stiffened and had paper pastedowns, both since removed. Underneath, the quires are sewn onto thongs. The leather thongs visible now were added in a repair in the 1960s and glued over older ones of some kind of fibre, which may be the original fifteenth-century sewing supports (Zandbergen, 1996–2026). The sewing holes in the folds of the leaves are worn, and in at least one quire there is a second set of holes that does not match the present binding, which is one of the pieces of evidence that the book has been taken apart and put back together.

IS IT GENUINE?

The question is still asked, so it is worth answering directly. Every physical test applied to the manuscript has come back consistent with an early-fifteenth-century European book: the parchment, its date, the ink, the pigments, the quire signatures, the wear, the sewing. A forger working before 1912 would have needed to acquire well over a hundred leaves of blank fifteenth-century calfskin, including several of the oversize sheets, and would have needed to anticipate protein analysis and accelerator radiocarbon dating. The alternative, that Voynich or a predecessor found a blank fifteenth-century book and filled it, founders on the same point: blank



The limp vellum cover, added in the eighteenth or nineteenth century. Beinecke Rare Book & Manuscript Library, Yale University.



Inside the front cover: traces of the removed paper pastedown. Beinecke Rare Book & Manuscript Library, Yale University.

medieval books of this size and construction do not survive, because blank parchment was reused.

What the materials cannot tell us is whether the text means anything. A real fifteenth-century book can contain nonsense as easily as sense. That question belongs to the text, and to Part II.

3

Where It Has Been

A book's history is usually written in it: owners sign their names, librarians add shelfmarks, readers scribble. This book has very little of that, and the little it has is faint. Its documented history begins nearly two centuries after it was made, in Prague, and for the century before that there is nothing at all. What follows is the record as it stands, with the gaps left open (Clemens, 2016; Zandbergen, 1996–2026, 2016).

PRAGUE

At the foot of the first page, invisible in ordinary light and only partly legible under ultraviolet, is a signature: *Jacobj à Tepenecé*, with what appears to be a shelfmark beneath it. Jakub Hořčický, a pharmacist and distiller who had made money and friends at the court of the Emperor Rudolf II, was ennobled in 1608 with the title "de Tepenec" and died in 1622. The signature therefore dates from between those years, and is the earliest certain fact about the book's ownership. Someone erased it, probably not long after; Wilfrid Voynich's attempt to revive it with a chemical reagent in the 1910s is the reason the corner of the page looks as it does (Davis, 2025).

The next witness is a letter. In 1639 Georg Baresch, an alchemist living in Prague, wrote to the Jesuit polymath Athanasius Kircher in Rome, who had just published a book claiming to read Egyptian hieroglyphs, and sent him a copy of some pages of a manuscript he could not read, hoping Kircher could (Zandbergen, 1996–2026). Baresch describes the book: it is full of plants unknown to Germans, pictures of stars, and "chemical secrets." Kircher's reply does not survive, but the book stayed with

Baresch until he died, when it passed to his friend Johannes Marcus Marci, physician and rector of the university in Prague.

Marci sent the book itself to Kircher in 1665 or 1666, with a covering letter that is now tucked inside its front cover. The letter is the source of nearly everything that has been said about the book's early history. Marci tells Kircher that a former owner, Rudolf II, had paid six hundred ducats for it, and that the Emperor had believed its author to be the thirteenth-century English friar Roger Bacon. Marci is careful to attribute the Bacon story to a third party, Raphael Mnishovsky, a courtier who had died in 1644, and to say that he himself suspends judgement. No record of the purchase has been found in the imperial accounts, and Mnishovsky is not a reliable witness; but Tepenec's signature puts the book in Rudolf's Prague, and the sum, though large, is not impossible for a book believed to be by Bacon (Zandbergen, 2016).

Whether Rudolf had the book from John Dee, the English magus who was in Bohemia in the 1580s, is a story with no evidence behind it beyond the fact that Dee owned Bacon manuscripts and was in the right place. The older claim that the folio numbers are in Dee's hand is not accepted by palaeographers who have examined them (Davis, 2025).

ROME

Kircher's library passed, at his death in 1680, to the Jesuit college in Rome, the Collegio Romano. The book is not mentioned in any surviving catalogue there, which is not surprising; it was a curiosity, not a text. When the Italian state confiscated church property in the 1870s, the Jesuits moved their more valuable books into the private libraries of individuals to keep them out of the state's hands. The manuscript went into the collection attributed to Petrus Beckx, the Superior General of the order, and from there to the Villa Mondragone at Frascati, outside Rome, which the Jesuits used as a college (Hunt, 2016).

LONDON, NEW YORK, NEW HAVEN

In 1912 the Jesuits, short of money, quietly sold a group of manuscripts from the Villa Mondragone to Wilfrid Voynich, a Polish-born rare-book dealer with shops in London and, later, New York. Voynich bought about thirty; this was one of them. The sale was supposed to be confidential, and Voynich obscured the book's origin for the rest of his life, which did nothing to help later researchers (Hunt, 2016).

Voynich was not a quiet owner. He showed the book widely, had photostats made and circulated them to anyone who might read it, and promoted the Bacon attribution as far as it would go, which was quite far: in 1921 a professor of philosophy at the University of Pennsylvania, William Romaine Newbold, announced a decipherment based on microscopic marks within the strokes, in which Bacon had described the Andromeda nebula and the structure of cells. The marks turned out to be the cracking of the ink. Newbold's reading was demolished by John Manly in 1931, and the episode set the pattern for the next century: a confident solution, wide public-

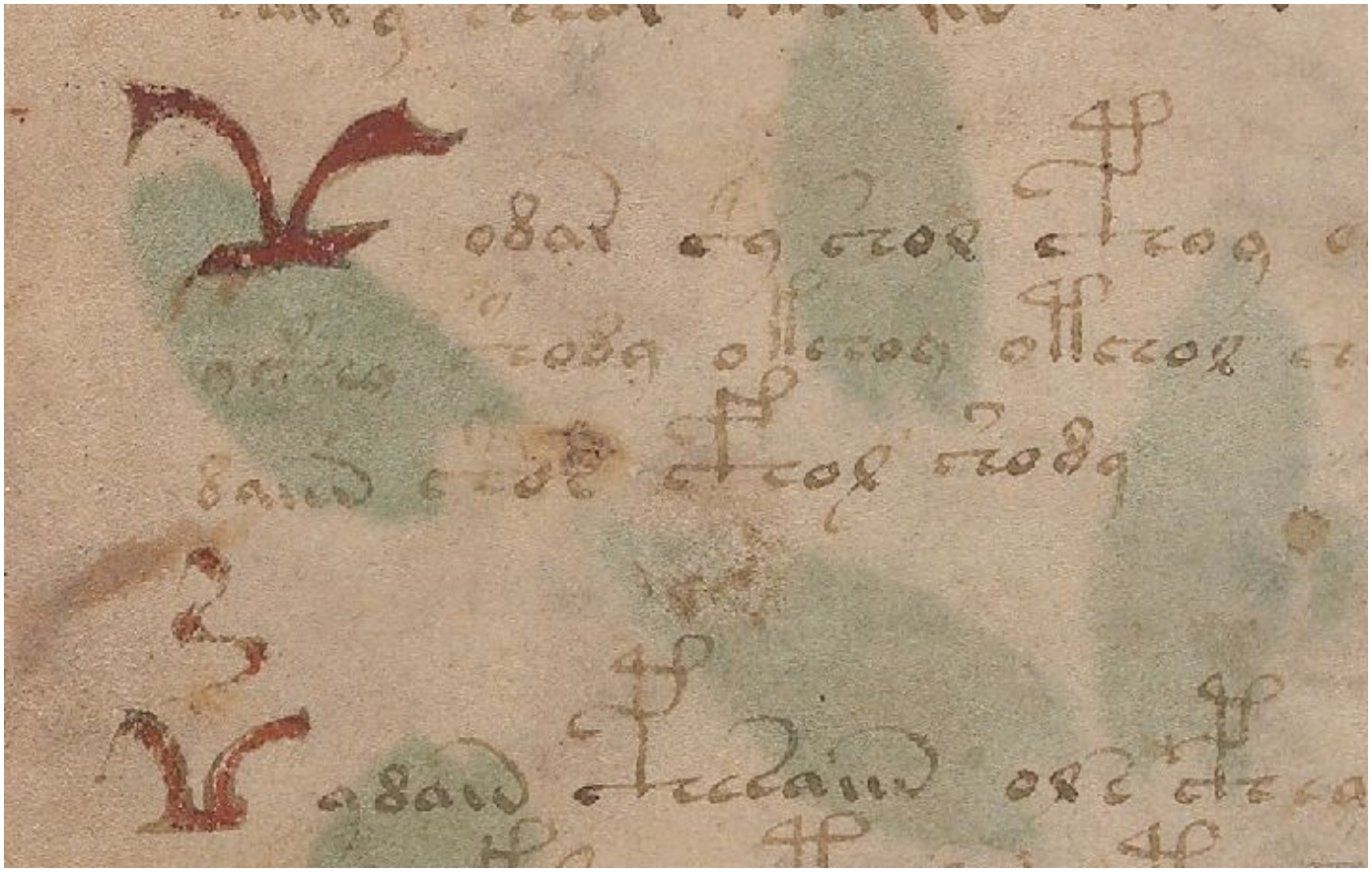
ity, and a collapse on inspection (D'Imperio, 1978).

Voynich died in 1930. The book passed to his widow, Ethel Lilian Voynich, the novelist, and on her death in 1960 to her companion Anne Nill, who sold it in 1961 to the New York dealer Hans P. Kraus for $24,500. Kraus priced it at $160,000, found no buyer, and in 1969 gave it to Yale University's Beinecke Rare Book and Manuscript Library, where it was catalogued as MS 408 (Clemens, 2016). It has been there since, photographed in full several times, and since 2020 the complete digital facsimile has been free to anyone with a web connection (Beinecke Rare Book and Manuscript Library, 2020).

WHAT THE RECORD SUPPORTS

Set out this way, the provenance is thin but coherent. The book was in Prague by about 1610, at the court where Rudolf II collected exactly this kind of thing. It passed through the hands of three men who tried to read it and could not. It spent two and a half centuries in Jesuit libraries, unread. It was bought by a dealer who tried to sell it as the work of a famous Englishman, failed, and left it to a library. At no point in that chain does anyone claim to understand it, and no one in the chain had the means to make it.

Everything before 1610 is inference from the object itself: the parchment date, the style of the drawings, the fifteenth-century quire numbers. On those grounds alone the book was made somewhere in central Europe, most likely in the German-speaking lands or northern Italy, in the first half of the fifteenth century. Where exactly, and by whom, the record does not say.



f. 1r, detail. The two paragraph-opening marks in red ink, unlike anything else in the manuscript, at the start of the second and third paragraphs. Beinecke Rare Book & Manuscript Library, Yale University.

4

In What Order?

A reader assumes that a book's pages are in the order its maker intended. For this book the assumption is probably wrong, and the evidence that it is wrong has been accumulating for twenty years. It reached a new stage in October 2026, when Colin Layfield and Lisa Fagin Davis published a statistical study arguing that the leaves were never meant to be bound as they are (Layfield and Davis, 2026). This chapter sets out what is known about the order of the book, which is more than one might expect, and what it means for reading it.

THREE LAYERS OF NUMBERING

There are three systems of numbers on the leaves, and they were added at different times by different people.

The quire marks come first. On the last leaf of each gathering, in the lower corner, someone wrote the number of the gathering as a Latin ordinal: *primus*, *2us*, *3us*, and so on to 20, with the "*-us*" written as a superscript 9, in the way fifteenth-century scribes abbreviated the ending. The numerals are in a hand and a style that palaeographers date to the fifteenth century, within a century of the writing of the book (Davis, 2025; Zandbergen, 1996–2026). Quire marks are added when a book is bound, so the book was bound in something like its present arrangement of gatherings while the century was still young. Quire marks 16 and 18 are missing because those gatherings are missing; marks 19 and 20 differ slightly in form from the rest, which Nick Pelling has taken to suggest more than one numberer (Pelling, 2006).

The folio numbers come second. They are writ-

ten in the upper right corner of each leaf in arabic numerals, in a hand that Davis places in the seventeenth century and that Yale's catalogue places in the sixteenth (Beinecke Rare Book and Manuscript Library, 2020; Davis, 2025). They run from 1 to 116 without resetting, which tells us two things: that the gatherings were in their present order when they were numbered, and that all 116 leaves were present. The fourteen leaves now missing were lost after the numbering, and in at least one case the loss is visible: a stub at folio 12 shows where a leaf was cut out (Davis, 2025). The fold-out leaves were numbered folded, so that their panels do not always carry the number where one would expect it, and the modern convention for naming the panels, f. 70v1, f. 70v2 and so on, was invented in the twentieth century.

The third system is modern: the panel names, and Yale's catalogue numbering of the digital images.

THE EVIDENCE THAT THE ORDER IS WRONG

What makes anyone doubt that the present order is the original one? René Zandbergen has collected the arguments over many years (Zandbergen, 1996–2026), and Davis has added physical evidence (Davis, 2025). The principal points are these.

Subject matter. The herbal pages are not all together. After fifty-six leaves of plants come the astronomical, zodiac, and bathing sections, and then, in quires 15 and 17, ten more leaves of plants that are not visibly different from the first fifty-six. The pharmaceutical pages are split between quires 15 and 19, and the style of the drawn containers suggests that the two leaves in quire 19 should come before the one in quire 15.

Scribes. In quires 4 to 7 of the herbal section, the work of Davis's Scribe 1 alternates with that of Scribe 2, sheet by sheet: the outer sheet of quire 4 (folios 25 and 32) is by Scribe 1, the next (26 and 31) by Scribe 2, the two inner sheets by Scribe 1 again. The same alternation produces the scattered language-B pages in figure 1.1. Davis, who has examined a great many medieval manuscripts, says she has never seen a book organised this way, and the simplest explanation is that sheets by the two scribes were once grouped and have since been shuffled.

A picture across the fold. On the back of folio 78 and the front of folio 81, in the bathing section, water drawn flowing from a vessel on the one page continues onto the other. The two pages form one design, which can only have been seen when the sheet they belong to lay open at the centre of its gathering. It does not; another sheet is bound inside it (Zandbergen, 1996–2026).

A stain. A large water stain runs through the upper margins of the first several dozen leaves. If the leaves had been in their present order when the water struck, the stain would shrink or grow steadily from leaf to leaf. It does not: on folio 32v it is smaller than on the facing 33r. Davis concludes that the book got wet, was taken apart to dry, and was put back together in a different order, all before the seventeenth-century folio numbers, which sit on top of the stains, were added (Davis, 2025).

Sewing holes. In quire 9, a single large fold-out sheet, there is a set of sewing holes in a fold other than the one now in the spine. If the sheet had been bound at that fold, its quire mark would sit where

quire marks normally sit. The sheet was refolded and resewn (Davis, 2025; Zandbergen, 1996–2026).

Paint. As noted in chapter 2, some of the clumsier paint has offset onto facing pages, and the offsets match the pages that face them now. The careless painting therefore happened after the reordering. Whoever coloured those plants did not know what order the book should be in, which means, as Zandbergen points out, that the colours are not a reliable guide to the plants (Zandbergen, 1996–2026).

THE SINGULION HYPOTHESIS

Layfield and Davis's 2026 paper adds a statistical argument and a new proposal (Layfield and Davis, 2026). They applied latent semantic analysis, a method for measuring the similarity of texts by their vocabulary, to blocks of the transcribed text at the level of pages, leaves, sheets and quires, and ran the same tests on two ordinary medieval books, the Aberdeen Bestiary and a copy of the *Speculum humanae salvationis*, as controls. In the controls, adjacent pages resemble one another in the way one expects of continuous text. In the Voynich manuscript, some adjacent pages do not, and some pages far apart do.

Their proposal is that the book was conceived not as a set of nested gatherings but as a sequence of single folded sheets—"singulions" in the authors' term—each a unit in itself, to be read one after another. A binder who nested those sheets into quires without being able to read them would produce exactly the pattern observed: scribes alternating sheet by sheet, a picture split across a fold, a text whose near neighbours are not always its nearest relatives. The authors are explicit that this reorders the problem rather than solving it. It does not produce a translation, and the original sequence of the sheets has not yet been recovered; they propose to use the water stain, whose shape should be continuous in the true order, as a guide.

The paper is three days old as this book goes to press and has not yet been examined by other researchers. Some of its conclusions rest on measurements of the text that depend, as all such measurements do, on the transcription and on choices about what counts as a unit. The physical arguments do not depend on those choices, and they were already strong.

WHAT FOLLOWS FOR A READER

Three things follow.

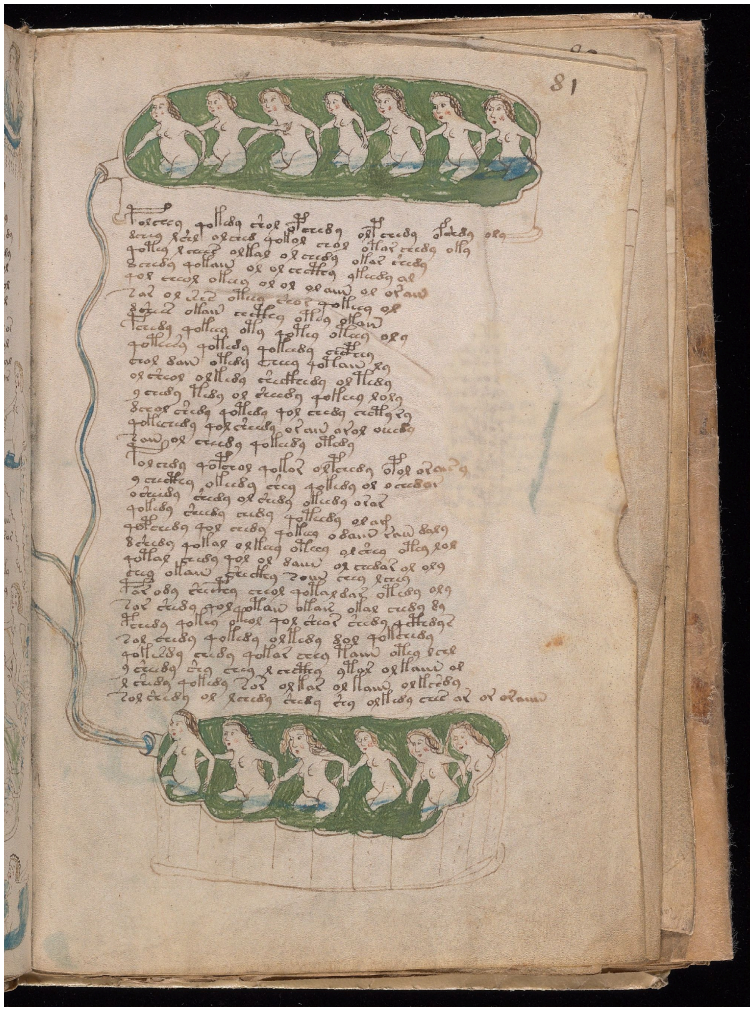
First, any analysis of the text that treats page order as meaningful—looking for a narrative, for a table of contents, for an index at the back, for a first chapter—is standing on sand. Analyses that treat the page, or better the sheet, as the unit are on firmer ground, and that is how the rest of this book proceeds.

Second, the sections are real but their boundaries are approximate. A leaf of plants in quire 17 is a herbal leaf; where it originally stood, we do not know.

Third, and more hopefully, the disorder is a kind of evidence. If the sheets can be put back in order by their physical traces, the result can be checked against the text: a reconstruction that also makes the text's vocabulary flow continuously from sheet



f. 78v: water flows from the vessel at the right edge… Beinecke Rare Book & Manuscript Library, Yale University.



…and continues on f. 81r. The two pages were drawn as one opening; another sheet is now bound between them. Beinecke Rare Book & Manuscript Library, Yale University.

to sheet would be a strong sign that the text means something and that the measurements are tracking it. That test has not been run. It is one of several we recommend in chapter 19.

Part II

The Script

5

Writing It Down

NOTHING can be counted until it has been written down in a form that can be counted, and for this manuscript that step is neither obvious nor finished. Every statistic in this book, and every statistic in the literature, rests on a transliteration: a decision about what the distinct symbols are, and a rendering of each page as a string of Latin letters standing for them. The decisions are not innocent. Some of the most-quoted "facts" about the text are facts about the transliteration.

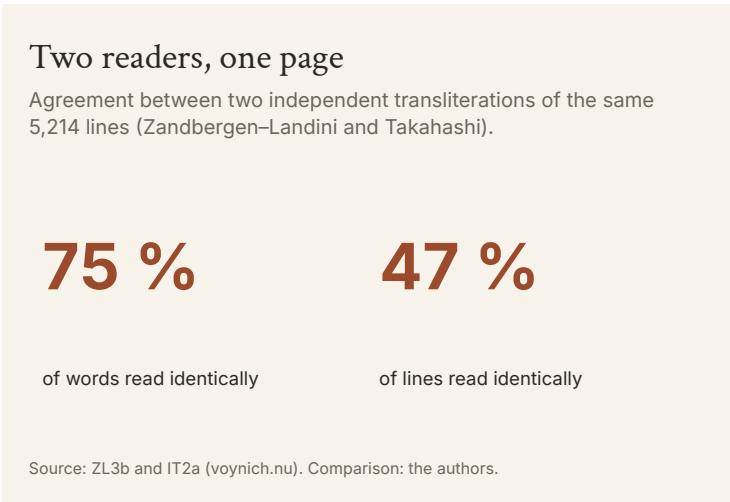


Figure 5.1 Two readers, one page.

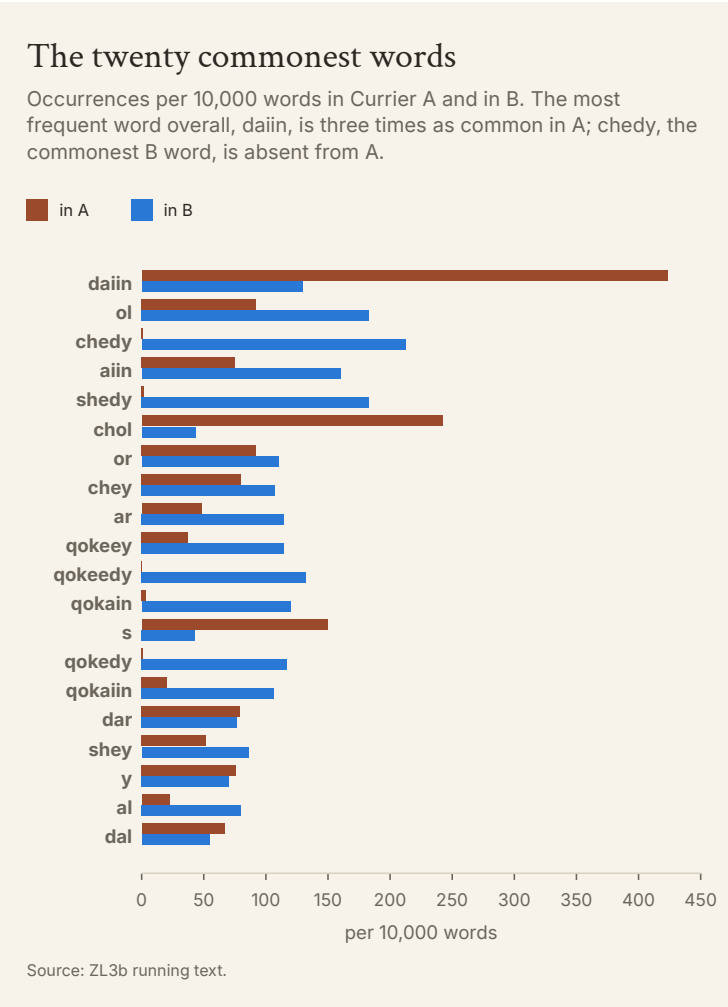


Figure 5.2 The twenty commonest words, by language.

WHAT THE SYMBOLS ARE

Look at a line of the text and the first impression is of an alphabet of modest size, perhaps twenty to thirty characters, written with a few basic strokes. Several characters resemble Latin letters or the abbreviation marks of Latin manuscripts: a form like *a*, a form like *o*, a form like the numeral 9 used by scribes for *-us*, a form like 4, and a looped form that medieval scribes used for *-is* or *-ris*. Four tall characters with loops on top—the "gallows"—have no close parallel. A pair of forms like *c* joined by a bar across the top, which scholars call benches, can also be written with a gallows letter planted in the middle, producing composite forms.

The difficulty is deciding where one symbol ends and the next begins. Is the sequence that EVA writes **ch** one letter or two? The strokes are joined, and it behaves in the text as a unit. Is **iin** three strokes of one letter, like the minims of *m*, or a run of three? Is the gallows-in-a-bench one character or two superimposed? Davis, using the Archetype digital palaeography platform to compare every instance, concluded in 2020 that the inventory is small and that the scribes wrote the same forms the same way, which is what one would expect of a practised script and not of an improvised one (Davis, 2020).

FOUR ALPHABETS

Over the past seventy years four main transliteration alphabets have been used (Zandbergen, 1996–2026). The First Study Group under William Friedman produced one in the 1940s and 1950s. Prescott Currier produced his own in the 1970s, using capital letters and digits (*4OFCC89* for the word EVA writes **qokeedy**). In the 1990s Gabriel Landini and René Zandbergen designed the European Voynich Alphabet, EVA, which uses lower-case Latin letters chosen so that the transliterated text is pronounceable and easy to type; it has been the standard for machine-readable work since. Glen Claston's v101 alphabet, from the 2000s, splits more finely, distinguishing variant forms that EVA merges.

The alphabets differ in what they treat as a unit.

EVA writes the commonest bench as two letters, **ch**, and the three-stroke ending as **iin**; Currier wrote each as a single symbol. Because EVA writes some single glyphs as two or three letters, statistics computed on EVA text are not quite statistics of the script. Bowern and Lindemann make the point carefully: EVA "probably under-differentiates" some glyphs and over-differentiates others, and the choice affects measurements such as entropy (Bowern and Lindemann, 2021). Throughout this book we therefore give results on the EVA letters and, where it matters, on a "glyph-unit" version in which the benches **ch**, **sh** and the four gallows-benches are counted as single symbols. The conclusions do not change; the numbers move a little.

THE TRANSCRIPTIONS

A transliteration alphabet is only useful once someone has applied it to all 226 pages and fold-out panels, and that has been done several times, by different people, with different results. The Zandbergen–Landini transliteration, begun in 1999 and revised continuously since, is the most complete: it records every piece of writing on every page, including labels, circular text and marginalia, with the position of each, and a set of variables for each page recording its quire, its illustration type, its Currier language and its scribe. We use its version of May 2025, ZL3b, throughout (Zandbergen, 1996–2026). Takeshi Takahashi's transliteration of the late 1990s, made independently, is the other full one in EVA.

The two can be compared line by line, and the comparison is the best available measure of how much of the "text" is actually settled. On the 5,214 lines both transcribers covered, the two agree exactly on 47 per cent of lines. On a word-by-word alignment they agree on 75 per cent of words. The disagreements are mostly about spaces—whether a gap is wide enough to count as a word break—and about a handful of ambiguous letter forms, **a** against **o**, **r** against **s**, the number of strokes in an **i**-run. Both transliterations mark their own uncertain spaces and uncertain letters, and we have followed the Zandbergen–Landini markings, treating uncertain spaces as spaces and dropping the 140 words that contain an unreadable character.

The practical consequence is that any claim resting on the exact spelling of an individual word is weaker than it looks. Claims resting on thousands of words, where the transcribers' disagreements wash out, hold up, and those are the claims we make.

THE TEXT AS COUNTED

With those cautions, here is the text by the numbers. The ZL3b transliteration contains 5,225 separate items of writing: 3,885 lines of ordinary paragraph text, 299 labels on the zodiac pages, 194 labels on plant parts in the pharmaceutical section, 76 labels of stars, 63 of the small female figures in the bathing section, and the rest circular and radial text in diagrams and miscellaneous labels. Together they hold 37,691 words. The ordinary running text, which is the basis for most of what follows, contains 34,180 words after the unreadable ones are removed, of 7,466 distinct spellings. The 100 commonest spellings account for 44 per cent of the running

text and the 500 commonest for two thirds; 5,266 of the 7,466 distinct words occur only once.

The commonest word, **daiin**, occurs 764 times, 2.2 per cent of the running text. The next are **ol**, **chedy**, **aiin**, **shedy**, **chol**, **or**, **chey**, **ar** and **qokeey**. Figure 5.2 gives the twenty commonest and shows at once the single most important division in the text: the words that dominate the first half of the list, **daiin**, **chol**, **chor**, nearly vanish in some sections, and the words that replace them there, **chedy**, **shedy**, **qokeedy**, are almost unknown elsewhere. That division is Currier's two "languages," and it is the subject of chapter 8.

6

Every Letter Knows Its Place

In 1951 John Tiltman, a British cryptanalyst who had spent the war on Japanese military ciphers, was shown photostats of the manuscript by William Friedman and set to work on the words. What he found, and reported with some puzzlement, was that they were built like compounds from a kit. Each word seemed to consist of a beginning drawn from one small set of elements, a middle from another, and an ending from a third, and most letters belonged to only one of the three sets (Tiltman, 1967). He could not say what the parts meant. But the observation has been confirmed by everyone who has looked since, and it is the single most important structural fact about the script.

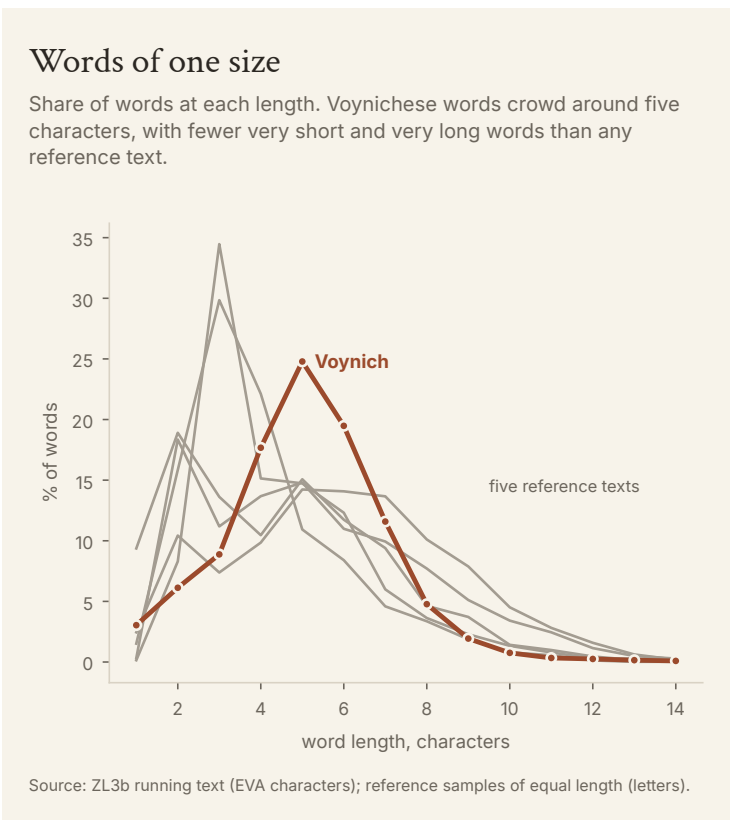
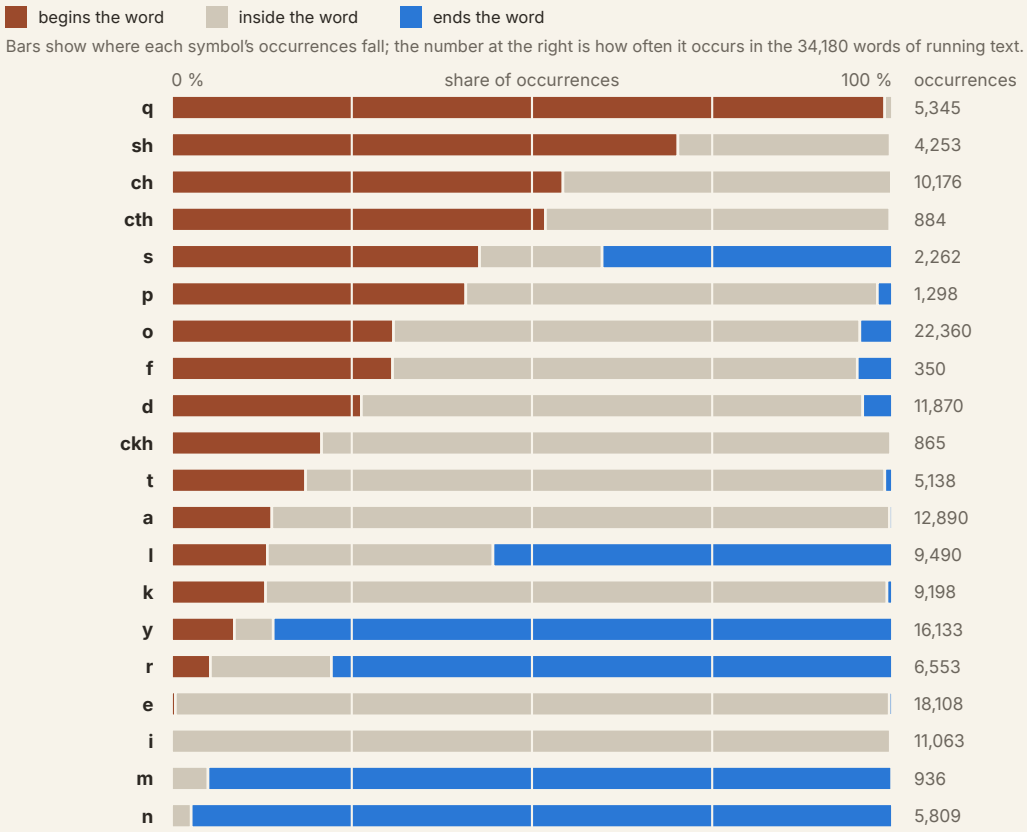


Figure 6.2 Word length, against five reference texts.

Every letter knows its place

In ordinary alphabets most letters can begin, end or sit inside a word. In Voynichese nearly every symbol has a fixed home. Any explanation of the text has to produce this pattern.



q is first in 99 % of its words, and 97 % of those are followed by o. No letter in a European alphabet behaves like this.

ch and sh open words; the 'bench' gallows cth, ckh, cph, cfh sit first or second, never last.

e, i and a are never first or last. They come in runs — ee, eee, ii, iii — found only mid-word.

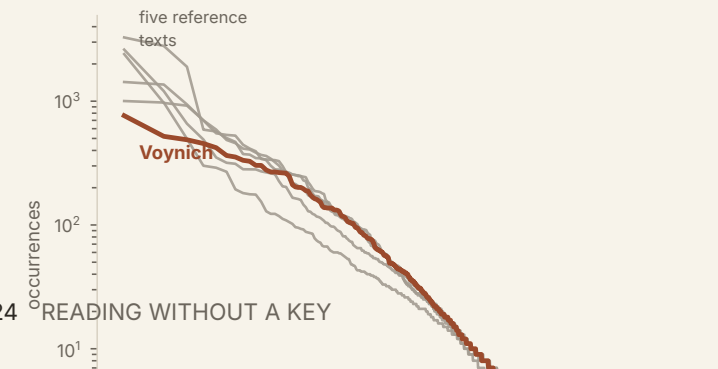
y ends 85 % of its words; n and m end 97 % and 94 %, and m clusters at the ends of lines.

Source: ZL3b transliteration, running text only; EVA alphabet with ch, sh, cth, ckh, cph, cfh counted as single symbols. Counts: the authors.

Figure 6.1 Where each symbol sits in the word.

Words obey Zipf's law

Rank of each word against how often it occurs, both on log scales. Texts in real languages fall on a line of slope about −1; so does Voynichese.



POSITION

Figure 6.1 shows the fact directly. For each of the twenty-two commonest symbols it shows how its occurrences divide between the start of a word, the inside, and the end. Three groups stand out.

At the top is **q**. It occurs 5,345 times in the running text, which makes it the twelfth commonest symbol, and 99 per cent of the time it is the first letter of its word. In 97 per cent of those cases the second letter is **o**. No letter of any European alphabet behaves like this; the nearest analogy is a letter that is really a prefix, or a mark that is written before a word rather than as part of it. Below **q** come the benches, **sh** and **ch**, and the gallows-benches **cth**, **ckh**, **cph**, **cfh**, which open a word two or three times out of four and never close one.

At the bottom are the closers. **n** ends 97 per cent of its words and **m** 94 per cent; neither ever begins one. **y** ends 85 per cent of its words and **r** three quarters. **l** is split, and is the one symbol that is at home anywhere.

In the middle are **e**, **i** and **a**, which are never first and never last. They occur in runs: **ee**, **eee**, **ii**, **iii**. The symbol **o** is the commonest in the text and is mostly a middle letter, but it begins a fifth of all words, usually before a gallows.

THE GALLOWS

The four tall letters **k**, **t**, **p**, **f** occur in just over half of all words, and almost never twice in the same word: 1.5 per cent of words contain two. They sit in a characteristic place, after the beginning and before the middle: **okaiin**, **qokeedy**, **otedy**, **choty**. Of the four, **k** and **t** are common, **p** and **f** rare, and **p** and **f** have a strong preference for the first line of a paragraph, which is the subject of the next chapter. Whatever the gallows are, a word has room for one of them, in one slot.

A GRAMMAR OF SLOTS

Tiltman's kit can be made precise. Jorge Stolfi, in the late 1990s, proposed a "core–mantle–crust" model in which every word has a core (typically a gallows or a bench), wrapped in a mantle of **e** and **o** and **a** letters, wrapped in a crust of the opening and closing letters (Stolfi, 1997–2000). There are several such models, and the differences between them matter less than what they share.

To make the point without choosing between them, we wrote the simplest version we could: an opening slot (one of **qo**, **o**, **d**, **s**, **ch**, **sh**, **y**, **a**, **l**, **r** or nothing), a gallows slot (one of the eight gallows forms or nothing), a middle slot (one of **ch**, **sh**, a run of **e**, **eo**, **o**, **a** or nothing), an optional single **d**, **l**, **r** or **s**, and a closing slot (one of **y**, **dy**, **ain**, **aiin**, **an**, **ar**, **al**, **or**, **ol**, **am**, **ey**, **eey**, **edy**, **eedy**, a single closing letter, or nothing). That template, which fits on three lines, matches 78 per cent of all word tokens in the running text, and a third of the distinct spellings.

What it fails to match is instructive. The commonest unmatched words are **dair**, **air**, **cheky**, **checkhy**, **olkeedy**, **olkeey**, **choty**, **olchedy**, **olkain**. Most of these are two template words run together: **olkeedy** is **ol** followed by **keedy**, **olchedy** is **ol** and **chedy**, **cheky** is **che** and **ky**. The frequent short word **ol** in particular attaches to the front of other words so often that

it behaves like a clitic. A more generous template that allows two cores covers over 90 per cent of tokens; the residue is mostly transcription noise and the long strings on labels.

LENGTH

A consequence of building words from a fixed number of slots is that they come out at nearly the same length. Figure 6.2 shows the distribution. Voynichese words have a mean length of 5.1 EVA letters and a standard deviation of 1.9; a quarter of all words are exactly five letters long. The reference texts, sampled to the same size, have broader distributions, and Latin's is much broader (standard deviation 2.7). What is striking is less the mean, which depends on the transliteration, than the shape: a tight, nearly symmetrical peak, with very few one- and two-letter words and a tail that falls off fast. Natural languages have a long tail of long words and a fat head of short function words. Voynichese has neither.

WHAT THIS DOES AND DOES NOT SHOW

These regularities establish that the text was produced by a procedure with rules, applied consistently across thirty-seven thousand words by five different scribes. That already excludes the lazier kinds of nonsense: a person scribbling freely does not keep **q** at the front of words 99 times in 100 for two hundred pages.

They do not, on their own, tell us what the rules are rules for. A slot structure of this kind can arise in several ways. It is what a syllabary looks like when written out: consonant, vowel, consonant. It is what an abbreviation system looks like. It is what a cipher looks like when each letter of the plaintext is replaced by a group of symbols drawn from a fixed table, a so-called verbose cipher, in which the plaintext letters show through as the slots (Bowern and Lindemann, 2021). And it is what a well-made generating procedure for meaningless text would produce, if its designer wanted the result to look like a language (Timm and Schinner, 2020). Telling these apart requires the rest of the evidence, and chapter 16 takes it up.

Two things can be said now. The first is that the slots are not arbitrary: they recur in exactly the places where a writing system would put them, and they interact with the line and the paragraph in ways that a generating procedure would have had to be designed to produce. The second is that the glyphs with the most fixed positions, **q** at the start and **n**, **m**, **y** at the end, are precisely the ones that are rare or absent on labels, where there is no line to be at the start or end of. That is the subject of chapter 7.

7

The Line and the Paragraph

In most written texts the line is an accident. A scribe writes until he reaches the margin and continues on the next line; the words do not know where the margin is. Prescott Currier noticed in the 1970s that in this manuscript they seem to (Currier, 1976). Words at the start of a line are spelled differently from words in the middle; words at the end are spelled differently again. The effect is large enough to see without statistics once one knows to look, and it is one of the clearest signs that the layout on the page is part of whatever system produced the text.

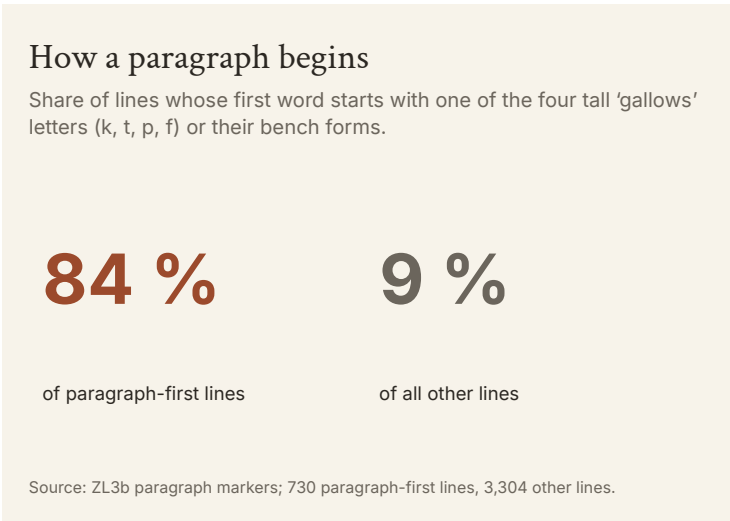


Figure 7.2 Paragraph openings.

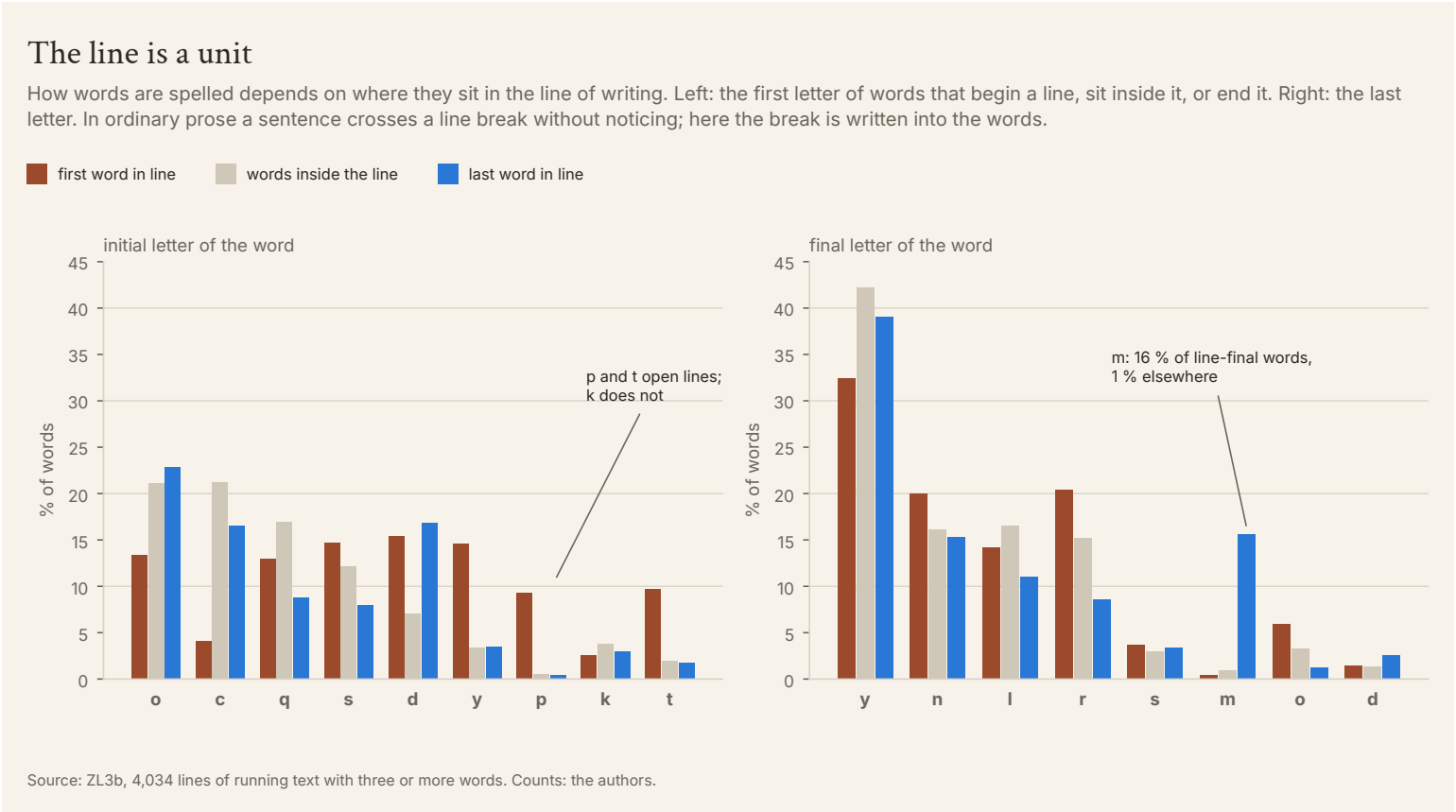


Figure 7.1 How spelling changes with position in the line.

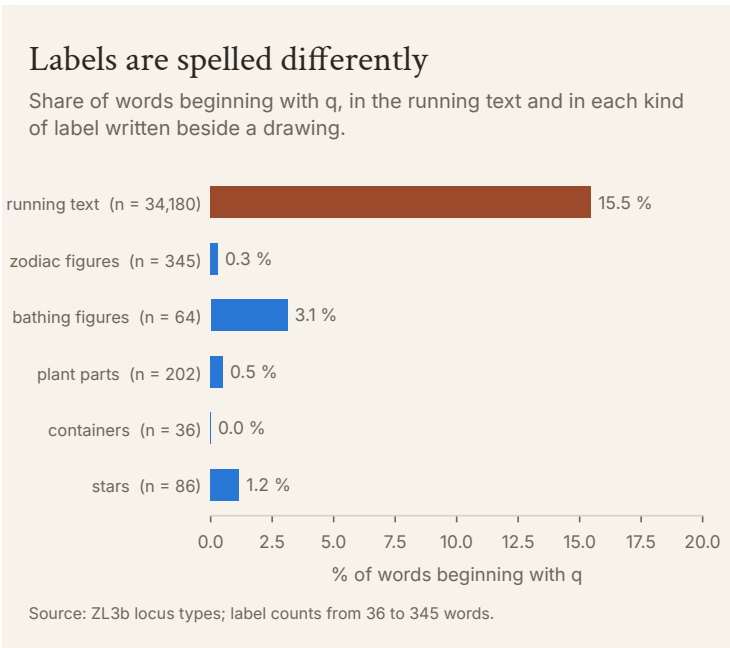


Figure 7.3 Labels against running text.

BEGINNINGS AND ENDS

Figure 7.1 shows the effect in the 4,034 lines of running text with at least three words. Take the first letter of each word. Among words that begin a line, 9 per cent start with **p** and 10 per cent with **t**; among words inside a line the figures are 0.5 and 2 per cent. Words beginning with **s**, **d** and **y** are also enriched at the line start, and words beginning with **ch**, the commonest opening inside a line, fall from 21 per cent to 4. Now take the last letter. Among words that end a line, 16 per cent end in **m**; among words inside a line, under 1 per cent. Words ending in **g**, rare everywhere, occur almost only at line ends.

The words themselves change length. The first word of a line averages 5.45 letters, a middle word 5.08, the last word 4.88. The pattern is consistent across sections and scribes.

PARAGRAPHS

The paragraph is a unit too. The Zandbergen–Landini transliteration marks the start of each paragraph; there are 736 of them. Of the 730 paragraph-opening lines with three or more words, 84 per cent begin with a word whose first letter is a gallows (**k**, **t**, **p**, **f**) or a gallows-bench. Of all other lines, 9 per cent do. Bowern and Lindemann report the same proportion, 85 per cent, from their own count (Bowern and Lindemann, 2021). The rare gallows **p** and **f** are concentrated there: 46 per cent of paragraph-opening lines begin with a **p**-word, against 1.3 per cent of other lines, and **f**, rare everywhere, is seventy times more likely at a paragraph opening than elsewhere.

Something similar happens at paragraph ends. On the first page, and on a few others (f. 8r, f. 25r, f. 42v, f. 46v), the final line of a paragraph is a single word or short phrase set off to the right, the kind of thing the transliteration calls a "title" and that Zandbergen has compared to the attribution at the end of a quotation. Elsewhere the last line of a paragraph is simply shorter, and ends, as we have seen, with a different set of letters.

WHAT KIND OF THING IS THIS?

There are two broad ways to read the line effects, and the literature has not settled between them.

One is that they are orthographic: a convention of writing, not of language. A scribe who writes **m** where he would otherwise write **iin** when he reaches the end of a line, or who adds a flourish to the first letter of a paragraph, is doing what medieval scribes did with Latin, where line-final abbreviations and enlarged initials are routine. On this view the words at line ends are the same words as elsewhere, in different dress. Bowern and Lindemann lean this way, calling the patterns "typographical rather than linguistic" (Bowern and Lindemann, 2021). Some evidence supports it directly: the words that begin lines with **s** look like ordinary words that begin with **o** or **a** with an **s** in front—**saiin** for **aiin**, **sol** for **ol**—as though the **s** were a line-initial mark rather than a letter.

The other reading is that the line is a unit of content: that each line is, or encodes, something complete, so that its first word and last word are a different kind of thing from its middle words, as the first and last words of a list or a verse would be. On this view the effects are real structure in whatever the text is.

The two readings are not exclusive, and the data do not yet decide between them. What the data do establish is that the text was not produced by writing out a continuous stream and breaking it at the margin. Either the writer composed line by line, or he transformed what he wrote at the line boundaries, and either way the layout was in his hands as he wrote. This matters for the explanations considered in chapter 16: a cipher that enciphered a Latin text letter by letter would not know where the lines were, unless the plaintext had already been laid out in the

same lines, which is possible but adds a step; a procedure for generating meaningless text would have to include a rule for line ends, which Timm and Schinner's does (Timm and Schinner, 2020).

LABELS

The strongest evidence that the line effects are about position comes from the words that have no line: the labels written next to drawings. The zodiac pages carry 345 label words, each beside a small figure; the pharmaceutical pages carry 202 beside drawings of plant parts and 36 beside containers; the bathing pages carry 64 beside figures and 55 beside tubes and vessels; 86 stand beside stars. Labels differ from running text in several ways (figure 7.3). Fewer than 3 per cent begin with **q**, against 15.5 per cent of running-text words; over half begin with **o**, against 20 per cent. They are a little longer on average, and nearly all unique: of 281 distinct zodiac labels, only 37 occur on more than one zodiac page, and the container labels, with a mean length of seven letters, occur once each and almost never in the running text.

Yet labels are made of the same parts. Half to four fifths of label words, depending on the type, occur as words in the running text as well; the zodiac labels **otaly**, **otal**, **okeey**, **oky** are ordinary vocabulary. The labels are, in short, words from the same system, stripped of whatever **q** does and of the line-initial and line-final marks, which is what one would expect of words standing alone.

Whether a label names the thing it stands beside is another question. In a herbal, the obvious expectation is that a label on a root is the name of the plant or the part; in a zodiac, that the label beside a figure is the name of a star or a degree. The expectation is reasonable and unproven. Attempts to use it as a key—to read **otaly** as a star name in some language—have so far produced readings that do not survive transfer to the next label, let alone to the running text (Bowern and Lindemann, 2021). The labels are a resource for a decipherment, and we return to them in chapter 17. They are not, by themselves, one.

8

Two Languages, Five Hands

Prescott Currier was a cryptanalyst, and cryptanalysts count. In the 1970s, working from the First Study Group's transcription, he counted letter pairs page by page and found that the pages fell into two groups so distinct that, in his words, they were "two different languages" (Currier, 1976). He called them A and B. He also observed that the two groups were written in two different hands, and that the hands and the languages went together. Both findings have held up for fifty years, and the second has been sharpened: in 2020 Lisa Fagin Davis showed that the hands are not two but five (Davis, 2020). The question this chapter addresses is what A and B are, and the answer, as far as the data go, is that they are not what Currier's word suggests.

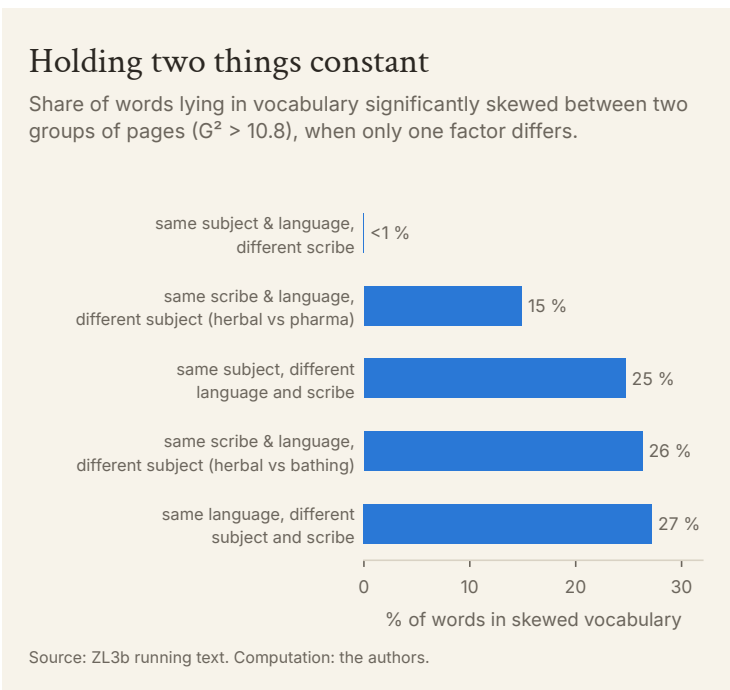


Figure 8.4 Controlled contrasts.

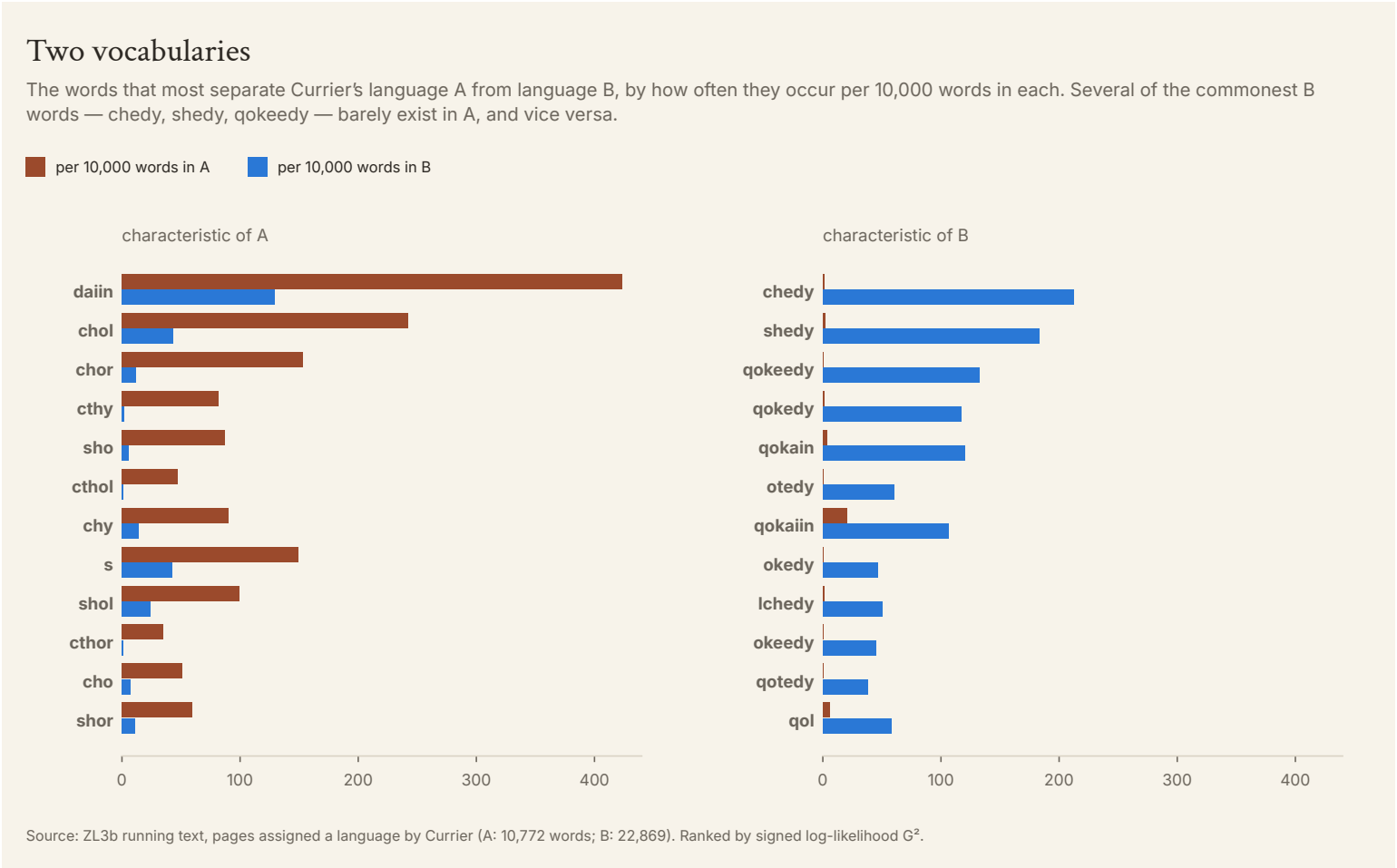


Figure 8.1 The words that separate A from B.

THE DIVISION

Figure 8.1 shows the words that most separate the two groups. Three of the ten commonest words in the manuscript, **chedy**, **shedy** and **qokeedy**, occur 486, 419 and 303 times on B pages and once, twice, and never on A pages. In the other direction **chol**, **chor** and **cthy** are common in A and rare in B, and **daiin**, the commonest word in the book, is three times as frequent in A as in B. The difference is not in a few words. It runs through the spelling: the letter pair **ed** occurs in 14 per cent of all words but is almost confined to B; words ending in **-edy** and **-eedy** are B words; words ending in **-ol** and **-or** lean to A.

Because the difference is in the spelling of common words, a score can be assigned to any page. Figure 8.2 does this for every page with at least thirty words of running text, in bound order, using a crude index: the share of words with B-type features minus the share with A-type features. The herbal section is mostly A, with B pages appearing as isolated spikes in quires 4 to 7 and again in quire 17. The bathing section is uniformly and strongly B. The recipe sec-

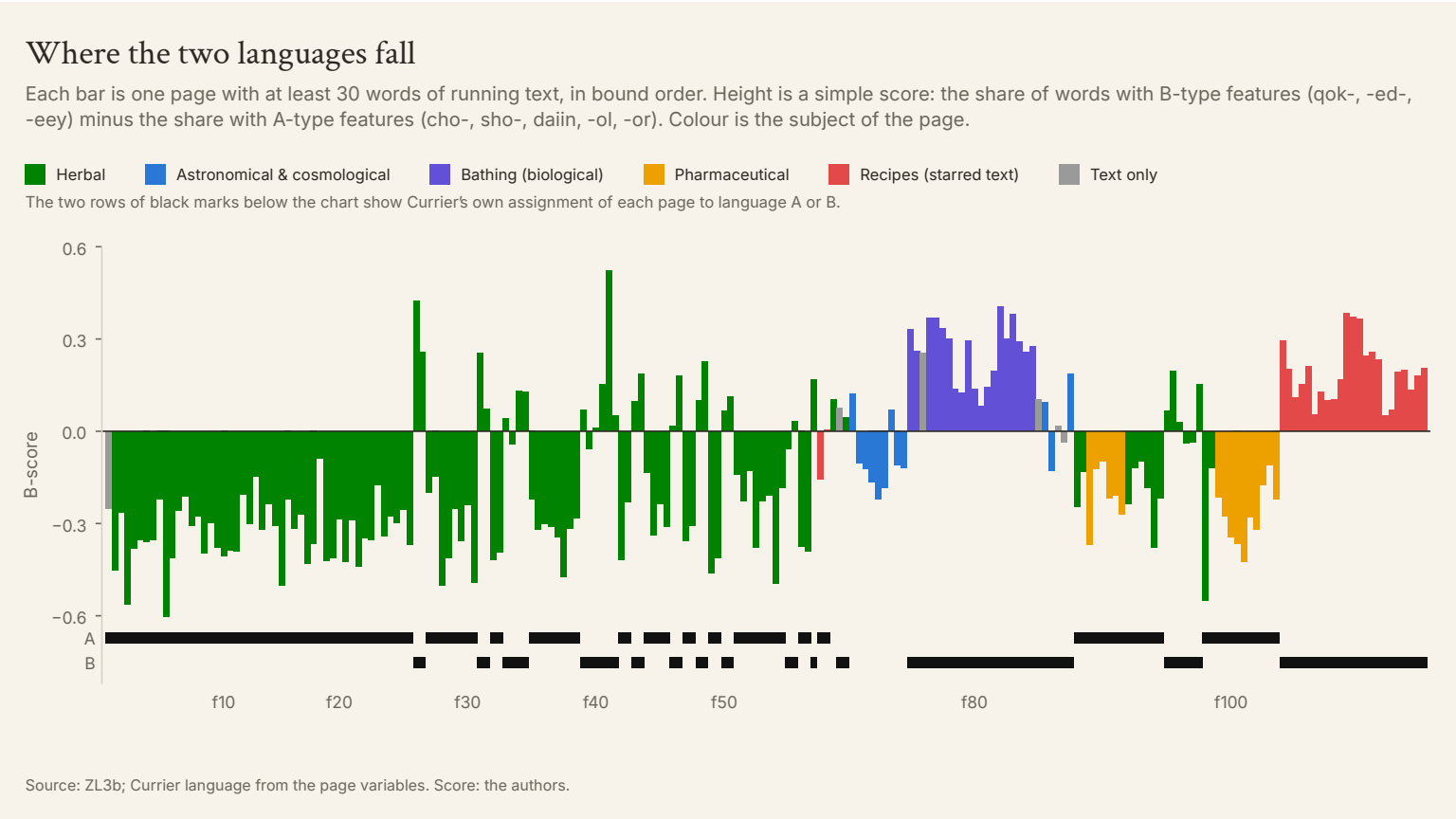


Figure 8.2 Every page scored for A or B, in bound order.

tion at the end is B, but less strongly. The pharmaceutical pages are A. The two small blocks of text pages, folios 58 and 85–86, are A and B respectively.

Two things are evident from the figure. The first is that A and B are not a continuum: pages are one or the other, and the crude score separates them nearly as cleanly as Currier's own assignments, shown as black marks beneath the bars. The second is that B is not one thing. The bathing pages and the recipe pages are both B, but their scores differ, and so does their vocabulary: **qokain**, **qol** and **shedy** dominate the bathing section, **aiin**, **ar**, **al** and **okeey** the recipes. Currier's two languages are better thought of as a family of registers, of which the herbal A and the bathing B are the extremes.

THE SCRIBES

Davis's five scribes map onto the languages but not onto the sections. Scribe 1 wrote the A pages of the herbal section and all of the pharmaceutical pages. Scribe 2 wrote the B pages scattered through the herbal quires, the whole of the bathing section, and the text pages of the rosettes sheet. Scribe 3 wrote the recipe section and the B herbal pages of quire 17, and also folio 58, a page of text with marginal stars which is in language A. Scribe 4 wrote the zodiac and

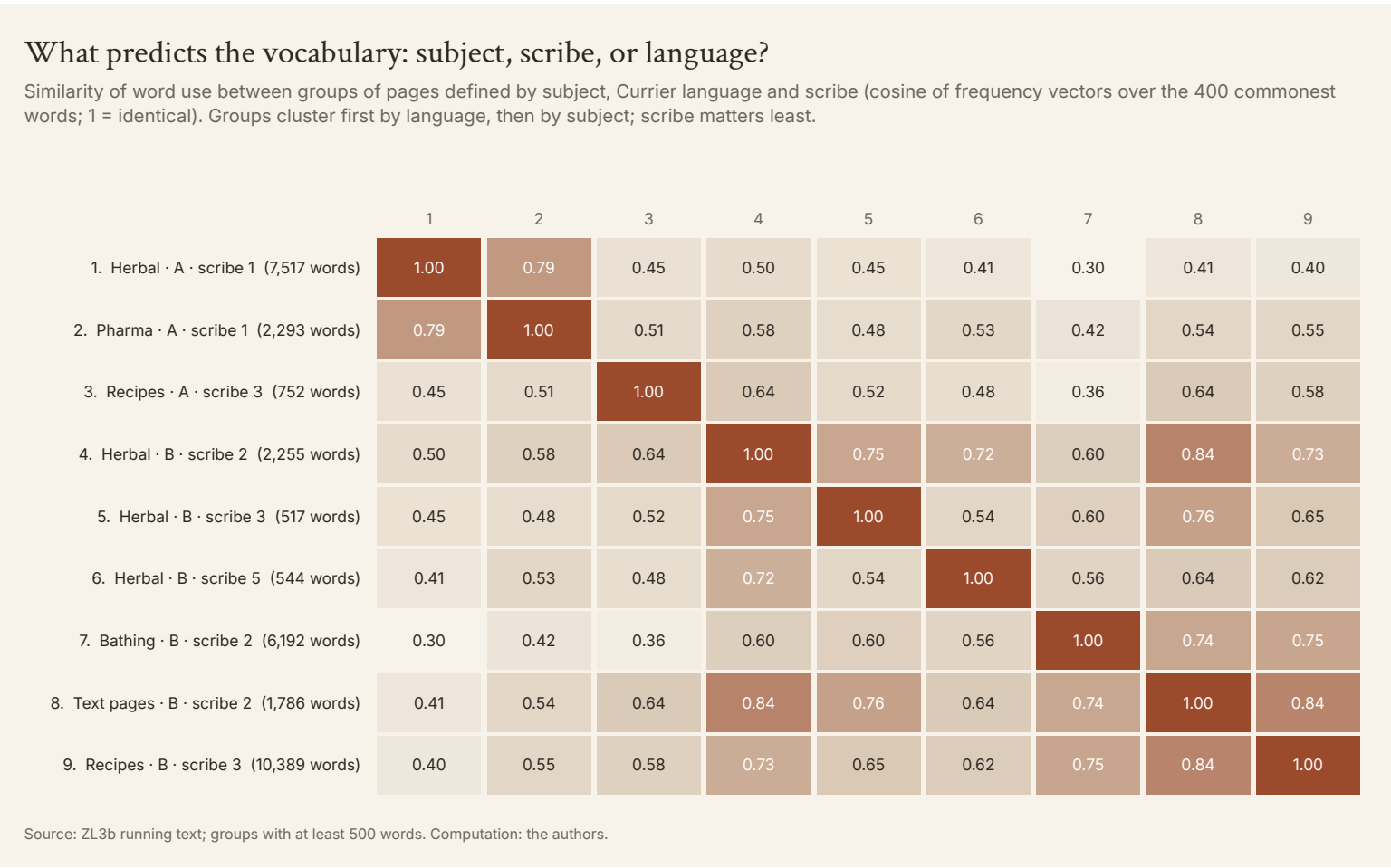


Figure 8.3 Vocabulary similarity between groups defined by subject, language and scribe.

astronomical pages, whose text is mostly labels and circles and was not assigned a language by Currier. Scribe 5 wrote a few herbal B pages in quires 6 and 8. So A is the work of scribes 1 and, once, 3; B is the work of scribes 2, 3 and 5 (Davis, 2020; Zandbergen, 1996–2026).

This raises the obvious question. Are A and B properties of the text, or just the habits of different writers? If Scribe 2 simply preferred **chedy** where Scribe 1 preferred **chol**, the "two languages" would be two handwritings, and would tell us nothing about content.

SEPARATING THE CAUSES

The manuscript allows a partial test, because its pages vary in three things at once—scribe, language, and subject—and for a few combinations two of the three are held constant while the third changes. Figure 8.4 shows the results. The measure is the share of words that lie in vocabulary significantly skewed between the two groups being compared.

Hold scribe and language constant and change the subject: Scribe 1's herbal pages against Scribe 1's pharmaceutical pages, both in A, differ in 15 per cent

of their words; Scribe 2's herbal pages against Scribe 2's bathing pages, both in B, differ in 26 per cent. The subject of a page, or something that travels with it, changes the vocabulary even when the same person is writing in the same language.

Hold subject and language constant and change the scribe: Scribe 2's herbal B pages against Scribe 5's herbal B pages differ in less than 1 per cent of their words. The sample is small, 544 words on Scribe 5's side, so the result is weak; but it is in the opposite direction from the handwriting hypothesis. Two scribes writing B on the same subject write the same words.

Hold subject constant and change both language and scribe: Scribe 1's herbal A against Scribe 2's herbal B differ in 25 per cent of their words, about as much as the two subjects within one scribe.

Figure 8.3 puts all nine groups large enough to compare on one grid. The groups cluster first by language—the three A groups together, the six B groups together—and within B by subject, with the bathing and recipe pages as the two poles and the herbal B pages between them. The scribe adds little once language and subject are known.

WHAT A AND B ARE

The pattern argues against the simplest explanations on both sides. A and B are not merely two handwritings, because two scribes writing the same language on the same subject agree, and one scribe writing two subjects does not. Nor are they merely two subjects, because the same subject, plants, exists in both, with the full A–B difference between its two halves.

What is left is that A and B are two ways of writing the same system that were in use at the same time in the same workshop, assigned to scribes and to sections together. This is consistent with a number of readings: two dialects of a language, two spelling conventions, two settings of a cipher, two stages in the development of a method. It is also consistent with Davis's suggestion, from the handwriting, that the manuscript was produced by a community of practice rather than a single author—several people who had learned the same system and used it slightly differently (Davis, 2025). Bowern and Lindemann report a related observation from the character level: the conditional entropies of A and B differ, 2.17 against 2.01 bits, but become nearly identical, 2.23 and 2.24, if the prefix **qo-** and the suffix **-dy** are removed (Bowern and Lindemann, 2021). Much of what makes B look different from A is the heavier use of two affixes.

Montemurro and Zanette showed in 2013, using an information-theoretic measure of how words cluster, that the manuscript's vocabulary is organised by section in the way that the vocabulary of a real book on several subjects is: words characteristic of the herbal section are rare elsewhere, and so on (Montemurro and Zanette, 2013). Our contrasts confirm this with the scribe controlled. Whatever the words are, they are distributed as though they were about something.

9

The Numbers

A Text that cannot be read can still be measured, and the measurements can be compared with those of texts that can. This chapter collects the measures that have proved most informative, computed afresh from the ZL3b transliteration and from five reference texts of exactly the same length: a stretch of the Latin Vulgate, a stretch of Pliny's *Natural History*, and passages of the Bible in English, German and Italian. The references are not meant to be candidates for the underlying language. They are there to show what ordinary text does, so that the manuscript's departures from it are visible.

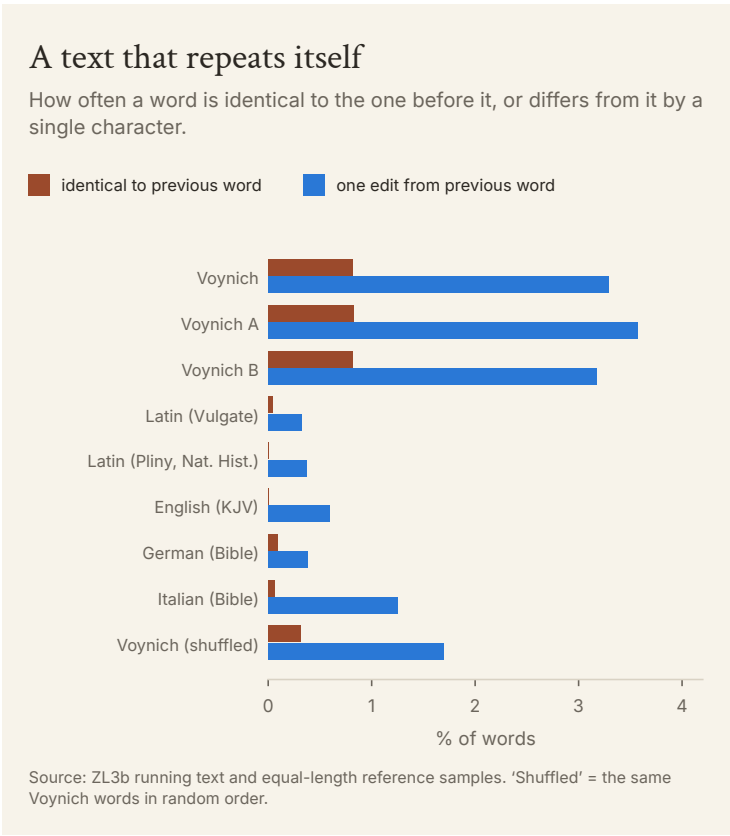


Figure 9.2 Repetition and near-repetition.

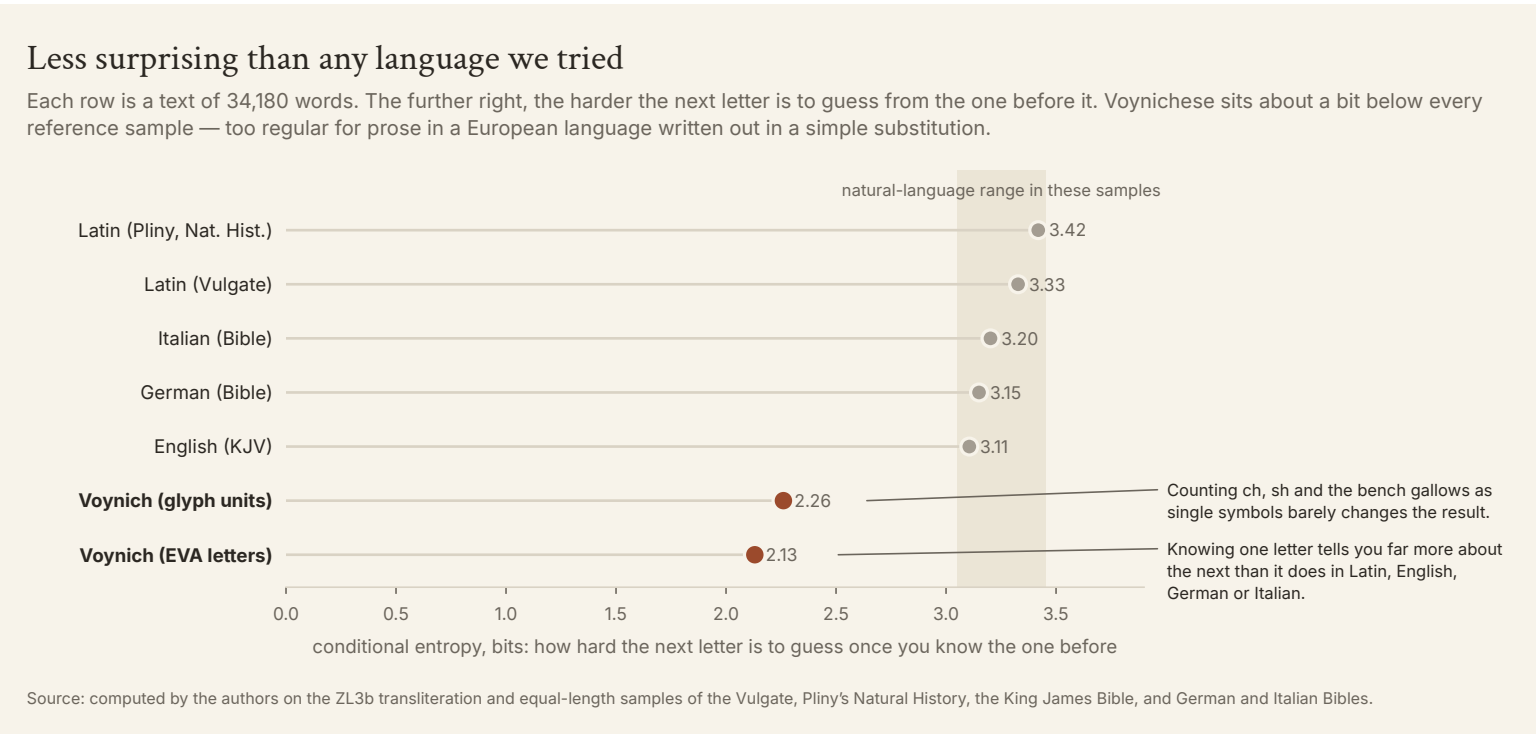


Figure 9.1 Conditional entropy against five reference texts.

ZIPF

In any large text the commonest word is about twice as frequent as the second, three times as frequent as the third, and so on: plotted on logarithmic axes, rank against frequency is a straight line of slope close to -1. This is Zipf's law, and it is true of every natural language, of most constructed languages, and, as it happens, of a good deal of nonsense. Voynichese obeys it (figure 6.3). The slope over the first thousand ranks is -1.04; the reference texts give -0.81 to -1.18. Gabriel Landini showed this in 2001 and it has been confirmed many times since (Landini, 2001). The law is a weak test: a text that failed it would be very strange, but passing it excludes little.

The shape of the curve carries more information than the slope. The manuscript's commonest word, **daiin**, accounts for 2.2 per cent of the running text. In the English sample, *the* accounts for 9.6 per cent, in the German *und* for 7 per cent, in the Latin *et* for 7.6. Voynichese has no word that dominates in the way a definite article or a conjunction does; its top of the list is flat. This is one of several signs that whatever the words are, the commonest of them are not function words of the European kind, or that function words are hidden inside other words.

VOCABULARY

In 34,180 words of running text there are 7,466 distinct spellings, and 5,266 of them occur once. The ratio of distinct words to total words, 0.22, is higher than the English sample (0.07) or the German (0.10) and close to the Latin Vulgate (0.20); only Pliny,

with his rich vocabulary of names and things, is higher (0.37). The hapax rate, 71 per cent of distinct spellings occurring once, is higher than any of the references except Pliny.

A high rate of unique spellings has two possible causes. One is a large vocabulary: an encyclopaedic text full of names. The other is spelling variation: many of the "distinct" words may be variants of one another. The second is certainly part of the story, because so many of the unique spellings differ from common words by a single letter. We return to this below.

ENTROPY

The most discussed statistic, and the one that most sharply separates the manuscript from ordinary writing, is the conditional entropy of its characters: how much the next letter is in doubt once the current one is known. William Bennett computed it in 1976 and found it far below that of any European language he tried (Bennett, 1976). Bowern and Lindemann repeated the comparison against 250 languages and found it below all of them (Bowern and Lindemann, 2021; Lindemann and Bowern, 2021).

Figure 9.1 shows our own version. The single-character entropy of Voynichese, 3.88 bits, is close to that of the references (4.00 to 4.06): the symbols are used about as unevenly as letters are. The conditional entropy is 2.13 bits. The references lie between 3.11 and 3.42. Treating the benches and gallows-benches as single symbols raises the figure to 2.26, which does not close the gap. In plain terms: in the Latin sample, knowing one letter cuts the uncertainty about the next by about 0.7 bits; in Voynichese, by 1.75 bits. The next letter is far more predictable than in any natural language written plainly.

Chapter 6 showed where this comes from. Most symbols have a fixed position in the word, so that once one knows the current symbol one knows roughly where in the word one is, and that restricts what can come next. A **q** is followed by **o**; an **i** by **i** or **n** or **r**; an **e** by **e**, **d**, **y** or **o**. The low entropy is the slot structure seen from a different angle.

What the entropy rules out is any simple letter-for-letter substitution of a European language. A text in Latin with each letter replaced by one symbol has Latin's conditional entropy exactly, 3.3 bits, not 2.1. Bowern and Lindemann tried the obvious manipulations—removing vowels, abbreviating, replacing letters by syllables—and found that none of them brings a natural text down to Voynichese levels; only "systematic conflation of distinctions," such as collapsing all vowels into one symbol, comes close (Bowern and Lindemann, 2021). Conversely a cipher that replaces each plaintext letter by a *group* of symbols with internal structure, a verbose cipher, lowers the entropy naturally, because the symbols within a group are predictable from each other. That is the direction in which this statistic points, and it is the direction Greshko's 2025 cipher takes (Greshko, 2025).

INDEX OF COINCIDENCE

The index of coincidence, the probability that two symbols chosen at random from a text are the same, is a tool of the cryptanalyst, used to distinguish simple substitution from polyalphabetic ciphers. For

Voynichese on the EVA letters it is 0.077; for the references, 0.070 to 0.074. The figure depends heavily on the transliteration and is not very diagnostic here, except in one respect: it is nowhere near the 0.04 of a text enciphered with a changing alphabet. Whatever the system is, it is consistent: the same thing is written the same way throughout.

REPETITION

The feature of the text that most struck its early readers, and that remains the hardest to explain, is how often it repeats itself. Figure 9.2 quantifies it. In the running text, 0.8 per cent of words are identical to the word immediately before them—**daiin daiin**, **chol chol**, **qokeedy qokeedy**. In the reference texts the figure is below 0.1 per cent. A further 3.3 per cent of words differ from the previous word by a single letter, against 0.3 to 1.2 per cent in the references. Within a window of five words, 16 per cent of Voynichese words have a near-twin, which is more than the Latin (4–7 per cent) and the German and Italian (10–14) and about the same as the English sample, whose repetitiveness is the King James Bible's.

The baseline matters. If the same 34,180 Voynichese words are shuffled into random order, the adjacent-repeat rate falls to 0.3 per cent and the near-repeat rate to 1.7 per cent. So the repetition is not merely a consequence of a small vocabulary; the text is locally far more repetitive than its own words in random order. Torsten Timm and Andreas Schinner built a theory of the manuscript on this observation: they propose that the text was generated by copying and slightly modifying words already on the page, a procedure they call auto-copying, which reproduces the near-repeat rate, the slot structure and the line effects without any underlying message (Timm and Schinner, 2020). Chapter 16 weighs that proposal. The observation itself is secure.

LONG-RANGE ORDER

The repetition is local. There is also structure at long range. Montemurro and Zanette computed, for each word, how much information its distribution across the text carries—whether it clusters in some parts and is absent from others, as the keywords of a book on several subjects do—and found that the manuscript's words cluster in this way to about the degree that words in real books do, and that the clustering follows the pictorial sections (Montemurro and Zanette, 2013). Our controlled contrasts in chapter 8 are the same finding with the scribe held constant. Amancio and colleagues, using network measures of word co-occurrence, likewise placed the manuscript among meaningful texts rather than among the shuffled versions they used as controls (Amancio et al., 2013).

THE SUM

Set side by side, the statistics describe a text that is more regular than language at the level of the symbol and about as regular as language at the level of the word and the section, with a layer of local repetition in between that language does not have. Any explanation has to produce all three. Simple substitution of a European language fails at the first. Random generation fails at the third. The candidates that sur-

vive are the ones in which the symbols are not letters: a verbose or homophonic cipher, a system of abbreviation, a syllabary or similar script for an unknown tongue, or a very well-designed generator. Those are the subject of Part IV.

Part III

The Pictures

10

Plants

ABOUT half the book is plants: 129 pages in the ZL3b count, nearly all in the first seven quires, with a scattering later (Zandbergen, 1996–2026). The format is as regular as a field guide. One plant to a page, drawn large, usually with root, stem, leaves and flower all shown; a block of text above the plant or beside it, sometimes continuing below; occasionally two plants, and occasionally a second block of text. The drawing was made first, in ink, and the text was fitted around it, which is the order in which illustrated herbals were normally made. The text on a herbal page averages 84 words. Some pages have a few lines, one has 163 words, and a handful have no text at all.

THE GENRE

The illustrated herbal is one of the oldest kinds of book there is, and by the fifteenth century it had a long tradition behind it, descending through Latin translations of Dioscorides and the pseudo-Apuleius herbal to the *Tractatus de herbis* family of the fourteenth and fifteenth centuries, in which each plant has its picture and a paragraph of properties and uses (Clemens, 2016). Anyone who has looked at a *Tractatus de herbis* manuscript recognises the Voynich herbal pages instantly as the same kind of thing: the same one-plant-per-page layout, the same emphasis on the root, the same habit of drawing the plant as a diagram of its parts rather than as it grows. Historians of botany who have examined it place it in that tradition while noting that it copies no known exemplar (Clemens, 2016; Zandbergen, 1996–2026).

That last point is the difficulty. In most herbals the pictures were copied from earlier herbals, and a plant can be identified by tracing its picture back through the copies to one that was labelled. The Voynich plants have mostly resisted this. A few look like real plants—there is a plausible water lily, several that could be members of the daisy family, something that looks like a viola—and one, on folio 35v, has a close match in a Paris manuscript of the *Tractatus de herbis* (Bibliothèque nationale, lat. 6823, folio 60r), noticed by several researchers (Zandbergen, 1996–2026). Many others combine the root of one plant with the leaves of another and the flower of a third, or have roots drawn as if they were animals or faces. Theodore Petersen and Ethel Voynich, in the 1930s, each compiled lists of tentative identifications; the lists agree with each other more than with later botanists, and with recent amateur identifications in some cases, but no list has won general acceptance (Zandbergen, 1996–2026).

Two cautions apply to all such attempts. The first is the paint. As chapter 4 showed, the colours were at least partly applied after the pages had been shuffled, by someone who did not keep inside the lines, so a blue flower may be blue because the painter had blue. The second is that the identification of a plant is only useful if it leads somewhere: if the label or the text beside it can be shown to contain the plant's name in some language. No identification has yet done that.

THE SUNFLOWER

The best-known identification is the worst. In 1944 the botanist Hugh O'Neill published a note identifying the plant on folio 93r as a sunflower and another as a capsicum pepper, both American plants, and concluded that the manuscript must postdate 1493 (O'Neill, 1944). The claim was widely repeated for sixty years. The radiocarbon date of 1404–1438 ends it: whatever the plant on folio 93r is, it was drawn before Columbus sailed. The episode is instructive, because O'Neill was a competent botanist and his sunflower looks, to a modern eye, like a sunflower. The drawings are suggestive enough to support almost any identification, and that is exactly why identifications without independent support do not count.

LABELS AND TEXT

Unlike the pharmaceutical pages, the herbal pages carry almost no labels: five words in the whole section, on three pages. Whatever names the plants have are in the paragraphs, not beside the drawings. The paragraphs themselves are the longest and most regular text in the book: 247 of the 736 paragraphs are here, and the paragraph-opening gallows rule of chapter 7 is at its strongest. In Currier's A pages the vocabulary is dominated by **daiin**, **chol**, **chor**, **shol**, **cthy**; the thirty-two B pages mixed among them use **chedy**, **shedy**, **qokedy**, as the bathing section does.

If the text describes the plants, the descriptions should have something in common with one another, and with no other section. Montemurro and Zanette's analysis found exactly this: the herbal section has its own characteristic vocabulary,

and the words that characterise it are among the strongest keywords in the manuscript (Montemurro and Zanette, 2013). Our contrast in chapter 8 confirmed that Scribe 1's herbal pages differ in vocabulary from his pharmaceutical pages. That is what a herbal text would do. It is also what a well-made imitation of a herbal text would do, if its maker varied the vocabulary section by section, and we do not know which we are looking at.

ORDERING

A real herbal is usually ordered: alphabetically by plant name, by the part of the body treated, by the plant's qualities, or following an older model's order. Order would give a way in; if the herbal were alphabetical, the first words or labels of successive pages would begin with the same letters in sequence. No such sequence is visible in the first words of the herbal paragraphs, which begin with the paragraph-initial gallows words, nor in the pattern of vocabulary across pages. Given that the sheets are out of order, this is less informative than it would otherwise be, and is one more reason to hope that the original sequence can be recovered.



f. 2r. A typical herbal page: a single plant drawn with root, stem, leaves and flowers, and a block of text above. Beinecke Rare Book & Manuscript Library, Yale University.



f. 35v: the plant with the closest known match in another herbal, Paris BnF lat. 6823, fol. 60r. Beinecke Rare Book & Manuscript Library, Yale University.



f. 41v: one of the thirty-two herbal pages in Currier's language B, written by Scribe 5. Beinecke Rare Book & Manuscript Library, Yale University.

11

Stars and the Zodiac

Quires 9 to 12 are the strangest part of the book to handle. Three of the four are single large sheets folded two, three or four times, so that opening them turns the small book into a broadsheet; one quire is a single ordinary bifolium. All of it is by one scribe, Davis's Scribe 4, who wrote nothing else, and almost none of it is paragraph text: of the 2,168 words in these quires, 276 are in paragraphs, the rest in labels and in rings of text running around circles. Currier did not assign these pages to either of his languages.

THE ASTRONOMICAL SHEETS

Quire 9 is one sheet with six panels, folios 67 and 68. On them are circular diagrams with a sun or moon face at the centre, rings of text, and scattered stars, each star with a label. One panel, folio 68r3, shows a cluster of seven stars with a label, which has been read since the 1920s as the Pleiades, the one asterism in the sky that a medieval reader would draw as seven dots in a bunch; next to it, a crescent moon. The identification is plausible and unprovable. Another panel, folio 68v3, has a circle divided into one half and two quarters, the T-O pattern that medieval cartography used to show the three continents (Zandbergen, 1996–2026). Quire 10, folios 69 and 70, holds further circular diagrams of the same family, with radiating text and stars.

Nothing in these diagrams has been matched to a known astronomical table or figure. There are no recognisable constellation outlines, no coordinates, no planets. The star labels, 86 words, are nearly all unique, which is what star names would be; but **dchol**, the one that recurs, is also an ordinary herbal

word, and more than half of the star labels occur in the running text.

THE ZODIAC

The zodiac occupies the rest of quire 10 and quires 11 and 12: twelve panels, folios 70v2 to 73v. Each is a wheel. At the centre is an emblem: two fishes, a ram, a bull, a pair of standing figures, a crayfish-like creature, a lion, a woman, a balance, a creature usually taken for a scorpion, an archer with a crossbow. Around the emblem run two or three rings of small female figures, most of them in tubs, each holding a star on a string, and beside each figure a label. The rings are interrupted by bands of circular text.

The sequence has three peculiarities, all noted by Zandbergen (Zandbergen, 1996–2026). It begins with Pisces rather than Aries, which is almost unknown in medieval zodiacs. Aries and Taurus each appear twice, on two panels with fifteen figures each, one with the animal drawn dark and one light, so that the two halves together make thirty; every other sign has its thirty figures on one panel (Pisces has twenty-nine). And Capricorn and Aquarius are missing, together with folio 74, which is exactly where they would have been.

Thirty figures to a sign is the number of degrees in a sign, and the obvious reading, which Zandbergen and most others accept, is that each figure stands for one degree (Zandbergen, 1996–2026). Medieval astrology had a literature of the degrees, the *paranatellonta*, in which each degree of each sign has its own image and meaning, descending from Hellenistic sources through Arabic intermediaries to Latin manuscripts of the kind that circulated in the fifteenth century. The Voynich wheels do not resemble any surviving degree-image series closely, and the figures are nearly interchangeable; what varies from figure to figure is the label.

THE MONTH NAMES

Here, and only here, the manuscript contains writing that can be read. In the centre of each zodiac panel, beside the emblem, someone has written the name of a month in Latin letters: March beside the fishes, April beside both rams, May beside both bulls, and so on to December beside the archer. The hand is not that of the text; the ink and the letter forms are later, and the forms are those of a Romance vernacular. The precise dialect has been argued over since the 1990s—French, Occitan, and northern Italian forms have all been proposed, and some of the spellings are unusual—and Zandbergen describes them simply as "a version of French" (Zandbergen, 1996–2026). The association of Pisces with March, Aries with April and so on is the ordinary medieval one, in which a sign is named for the month in which the sun enters it.

What the month names tell us is limited but real. Someone who could read the Latin alphabet and spoke a Romance language had the book, understood the emblems well enough to assign months to them, and did not add anything else: no plant names in the herbal, no title, no gloss on a single Voynichese word. They are the earliest readers we know of, and they could not read it either.

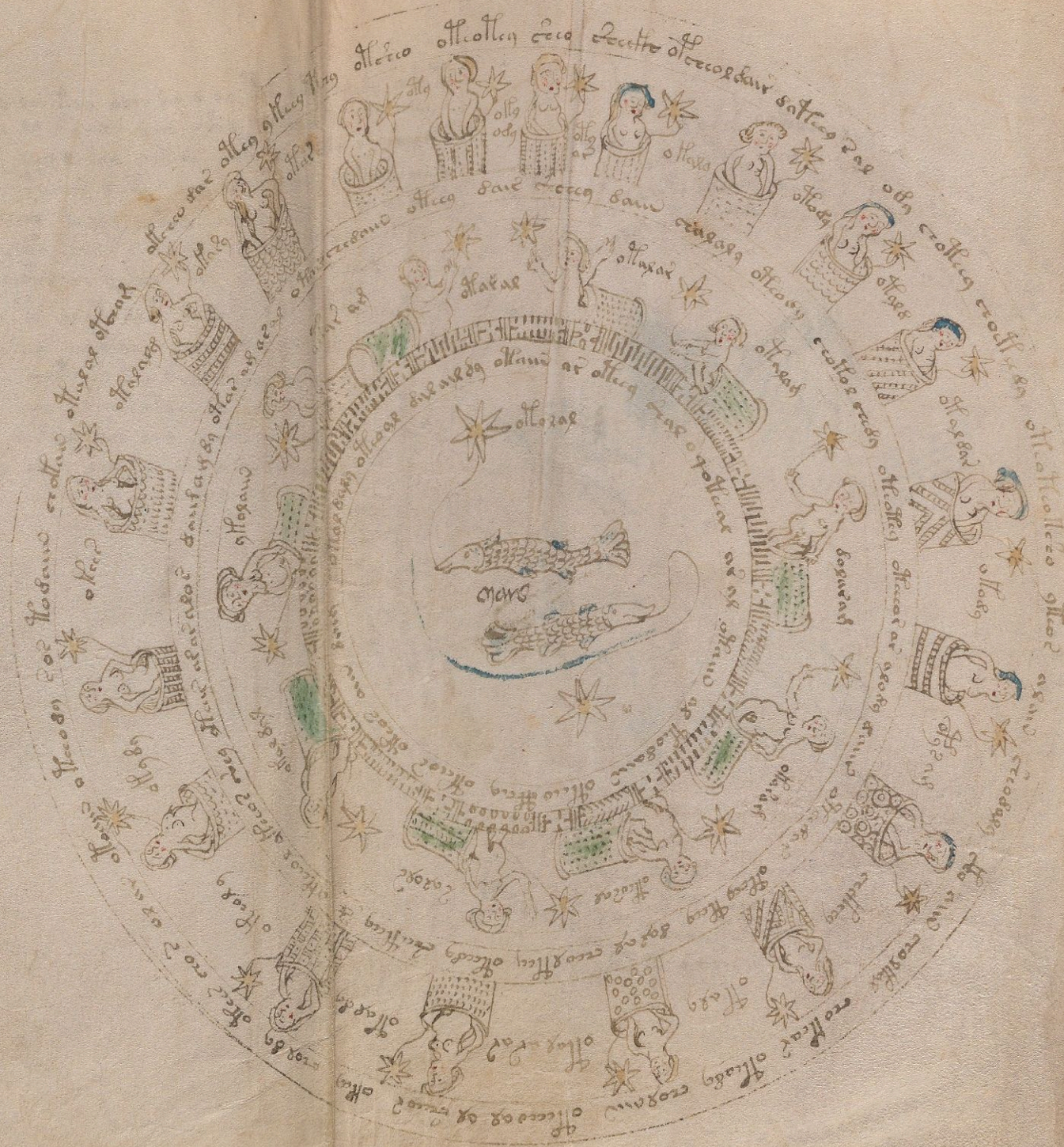
THE LABELS

The 345 zodiac labels are the largest set of labels in the book and the most tempting, because they stand beside figures that almost certainly represent something specific: a degree, or a star, or a day. If they were names, in whatever system the text uses, a decipherer would have 281 distinct names with known referents.

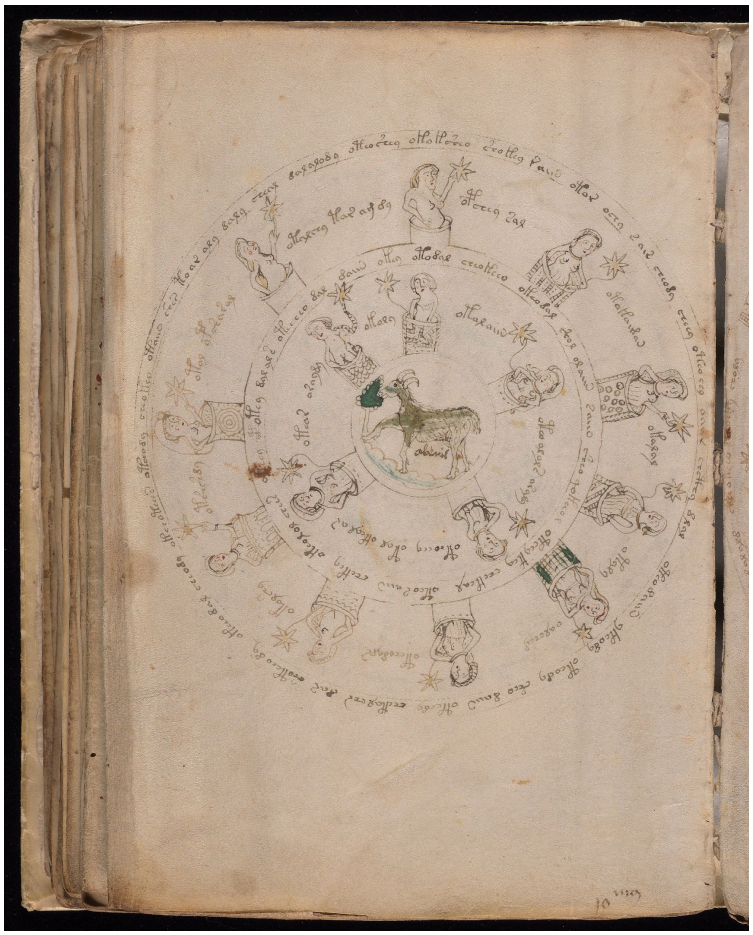
The distribution argues against the simplest version of this. Thirty-seven of the 281 distinct labels occur on more than one zodiac panel: **oky**, **ar**, **otaly**, **otal** and **okeey** each appear on four different signs. Names of degrees would not repeat across signs; names of stars would not either. And the labels are made of the ordinary vocabulary of the running text—over half occur there as words—rather than looking like a separate onomastic stock. Whatever they are, they are not thirty unique proper names per sign. They might be attributes of the figures, or abbreviated phrases, or something else; the evidence does not say.

COMPARANDA

The emblems themselves have been compared with zodiac cycles in hundreds of other manuscripts. The closest resemblances, in the crossbow of Sagittarius, the odd Scorpio, and the dress and headgear of the human figures, are with manuscripts from southern Germany of about 1400 to 1500 (Zandbergen, 1996–2026). This agrees with the radiocarbon date and with the one other datable feature of the drawings. It does not identify a source; no known zodiac has the Voynich arrangement of tubs and rings, and a full cycle beginning with Pisces has no known parallel.



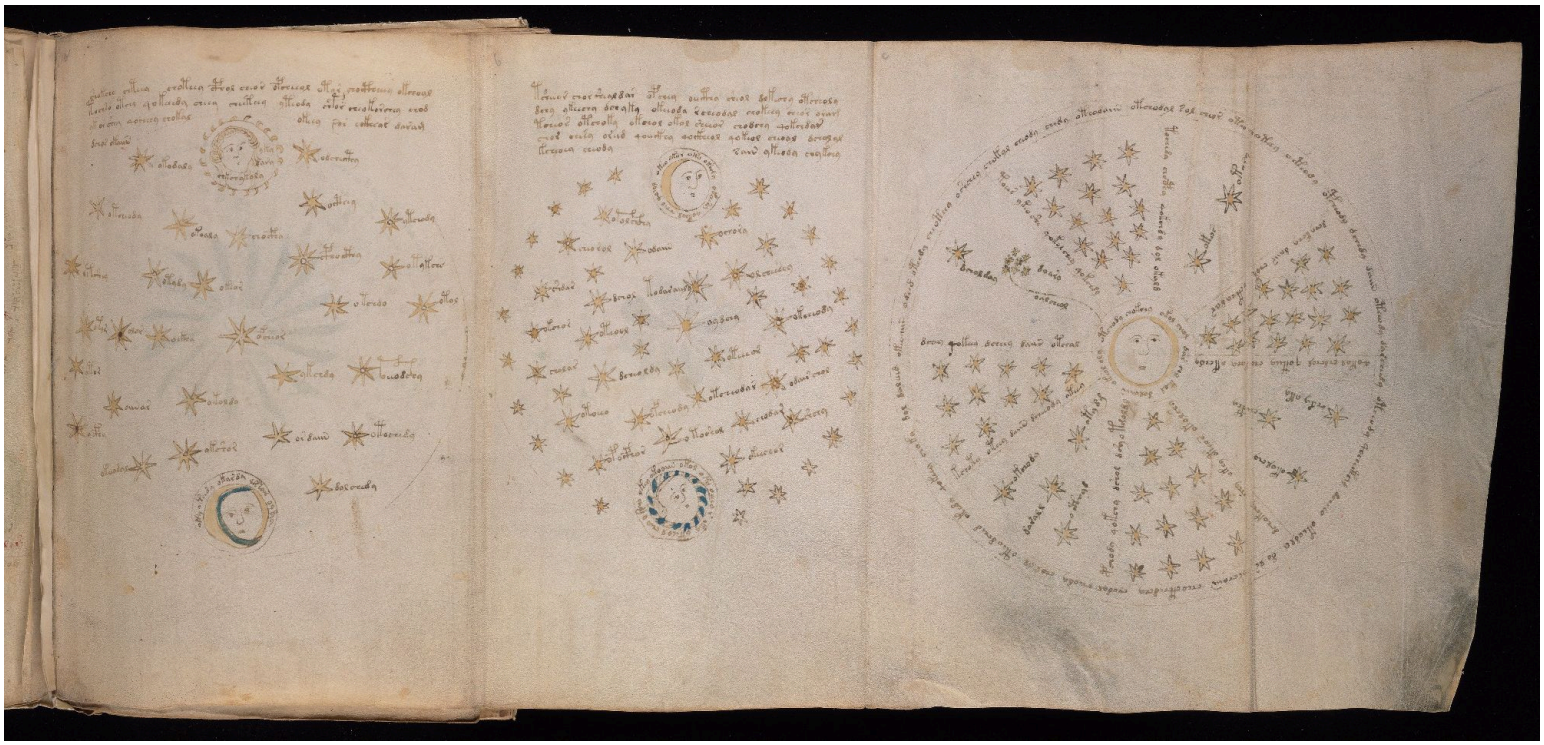
f. 70v2, Pisces: the first sign of the cycle, with twenty-nine figures in two rings and the month name written beside the fishes in a later hand. Beinecke Rare Book & Manuscript Library, Yale University.



f. 70v1, the first of the two Aries panels, with fifteen figures. Beinecke Rare Book & Manuscript Library, Yale University.



f. 73v, Sagittarius with a crossbow, the feature most often cited for a southern German origin. Beinecke Rare Book & Manuscript Library, Yale University.



f. 68r, the astronomical fold-out: rings of stars with labels, a sun and a moon, and at the right the cluster of seven stars often taken for the Pleiades. Beinecke Rare Book & Manuscript Library, Yale University.

12

The Bathers

Quire 13 is the most coherent gathering in the book. Twenty pages, folios 75 to 84; one scribe, Davis's Scribe 2; one language, the most extreme form of Currier's B; one subject. The subject is women in water. Page after page shows small, plump, mostly naked female figures standing in green pools, sitting in tubs, lying in channels, holding on to pipes. The pools are connected by tubes and conduits that run between paragraphs of text, so that the whole quire reads as a plumbing diagram with people in it. There are more than two hundred figures here, and over five hundred in the book as a whole (Zandbergen, 1996–2026). The section has been called biological, balneological, anatomical and gynaecological by different writers; we call it the bathing section, because that is what the pictures show, and leave what it is about open.

WHAT IS DRAWN

The figures are of a single type, the same type that holds the stars in the zodiac wheels: round-bodied, round-faced, with loose hair or a headdress, and a stylised posture. A few hold objects, among them rings and what may be flowers. The water is drawn in green and blue, sometimes with wavy lines for flow. The vessels range from plain basins to elaborate structures with crenellated rims and arches, and the pipes include spouts, funnels and what look like valves or taps. On folio 77r figures stand beneath a band of coloured arcs across the top of the page and down its left margin, each attached to a pipe; on folio 78r two pools, one above the other, are fed from a structure at the head of the page; on folio 84r a pool at the top is roofed with a row of figures and tubs, like a gallery.

Whether the pipes are pipes is a fair question. The drawings have been read as hydraulic engineering, as the vessels of the body with the figures as humours or spirits, as alchemical apparatus, and as the plumbing of a bathhouse. Zandbergen notes the resemblance of some of the objects to those in illustrated copies of the *De balneis Puteolanis*, a thirteenth-century poem on the healing baths near Naples whose manuscripts show bathers in pools and the equipment of the bath (Zandbergen, 1996–2026). Bathing as medicine—balneotherapy—was an established branch of medieval practice, and a treatise on the baths with pictures of bathers would not be out of place beside a herbal and a pharmacy. Beyond that resemblance of genre, no specific source has been found.

THE TEXT

The bathing section contains 6,192 words of running text in 86 paragraphs, more text per page than any other illustrated section, and it is where the B vocabulary is at its purest. **qokain**, **qol**, **qokeedy**, **shedy**, **qokal**, **shey**, **qokeey** and **qoky** are all far commoner here than anywhere else; **daiin**, **chol** and **chor** nearly vanish. The prefix **qo-** is at its most frequent. Chapter 8 showed that these differences hold even against the same scribe's herbal pages: whatever Scribe 2 was writing here, he wrote it differently from what he wrote beside the plants.

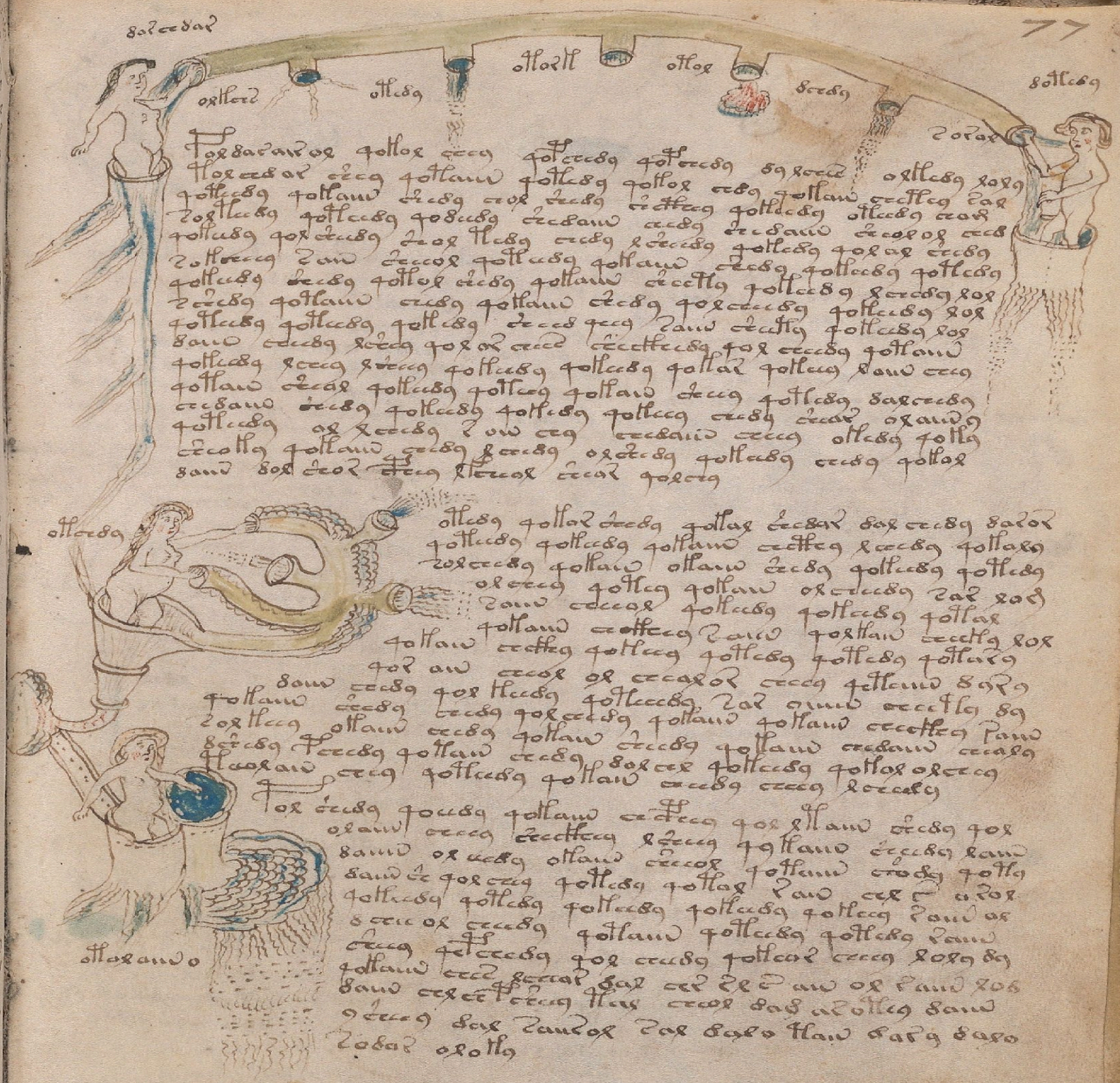
The 64 labels beside figures and 55 beside tubes and vessels follow the general rule for labels: rarely **q**-initial, mostly **o**-initial, and drawn from the ordinary vocabulary, of which four fifths of the figure labels are members. **otedy**, **otol**, **dal** and **daldy** each label more than one figure.

FOLIO 78 AND FOLIO 81

The section contains the single most useful piece of evidence about the book's order, described in chapter 4: a stream of water drawn from a vessel at the right edge of folio 78v continues from the left edge of folio 81r. The two pages belong to one sheet and were drawn as one opening. Another sheet, folios 79 and 80, is now bound inside them, so that the picture is broken. The sheet 78/81 was originally the innermost of its gathering (Davis, 2025; Zandbergen, 1996–2026). This is also the clearest case where the text might one day help: if the paragraphs on 78v and 81r are continuous, a reconstruction that joins them should show it, and Layfield and Davis's similarity measure is one way to look.

WHAT CAN BE SAID

Three things. The section is a planned unit: one sheet-maker, one scribe, one vocabulary, one iconography, drawn and written together. It is the part of the book whose text behaves least like the herbal text, which is some evidence that the text follows the pictures, that is, that it is about something. And it belongs to a recognisable medieval genre, the illustrated treatise on bathing and the body, without being a copy of any known example. The rest—what the women are doing, what the water is, why the pipes—is not something the manuscript tells us, and we do not tell the reader.



f. 77r. Figures along the top and left margins, each with a pipe; paragraphs of language-B text between. Beinecke Rare Book & Manuscript Library, Yale University.



f. 78r. Two pools, fed from the structure at the top of the page. Beinecke Rare Book & Manuscript Library, Yale University.



f. 84r. The last page of the section: a pool roofed by a gallery of figures. Beinecke Rare Book & Manuscript Library, Yale University.

13

The Rosettes

Quire 14 is a single sheet of parchment about 45 centimetres square, folded in six, and it is the only part of the manuscript that is famous as a picture. Unfolded, one side shows nine large circles arranged three by three, each filled with a different pattern—rays, scales, stars, tubes, a ring of tiny figures—and joined to its neighbours by causeways, pipes and scrolls. The other side carries smaller circular diagrams and three panels of plain text. The text is Currier B by Davis's Scribe 2, the scribe of the bathing section, which is one reason to think the sheet was bound near where it belongs (Davis, 2020).

WHAT IS DRAWN

Each of the nine circles is different. The central one has a ring of small towers or pavilions and a ring of small figures around a patterned core. Two circles, at opposite edges, have a starburst of blue rays. One is filled with a field of stars. One has, at its edge, a group of buildings with walls and towers whose battlements end in the forked merlons that are characteristic of northern Italian and Alpine fortification in the late Middle Ages—the detail most often cited in arguments for an Italian or South Tyrolean origin (Zandbergen, 1996–2026). In one corner of the sheet, outside the circles, is a small circle divided into a half and two quarters, the T-O schema of the medieval world map, which also occurs on folio 68v (Zandbergen, 1996–2026). Around and between the circles runs text: labels, rings of writing along the rims, and words along the connecting paths.

The sheet has been read as a map, a cosmological diagram, a plan of a city or a set of cities, an itinerary,

and a scheme of the elements. Most readings notice the same things: that the circles are connected in a way that implies movement or relation between them; that the central circle is the largest and most elaborate; that the castle makes one circle look like a place; that the stars and rays make others look like sky. None of these readings has been shown to be more than a resemblance.

ORIENTATION

Even which way up the sheet goes is uncertain. The folds were resewn at least once, the sheet was foliated folded, and the circles are drawn to be looked at from more than one direction, with text that runs around rims in both senses. We reproduce the sheet in the orientation of Yale's photograph, which is one of the possibilities, and describe positions on it without assuming that orientation is original.

THE TEXT

The rosettes sheet holds 1,678 words in the ZL3b count, of which 655 are in paragraphs, on three panels of the reverse, and the rest in rings and labels. The paragraph text is B of the bathing-section kind. Of all the parts of the book, this is the one where the words most obviously stand in relation to specific drawn things—a word beside a tower, a ring of words around a circle—and the one where that relation is least understood. It is, with the zodiac labels, the place a decipherment would be tested first.



f. 85v–f. 86r, the rosettes sheet unfolded, in the orientation of Yale's photograph. Nine circles joined by causeways; a walled town with forked battlements at the lower right; a small T-O circle at the lower left. Beinecke Rare Book & Manuscript Library, Yale University.

14

Jars and Roots

THE pharmaceutical pages are the part of the book that looks most like a working document. There are sixteen of them in the ZL3b classification, on fold-out leaves in quires 15 and 19, folios 88, 89, 99, 100, 101 and 102. Each shows rows of small drawings of plant parts—roots mostly, with some leaves and the occasional whole plant—laid out across the page like specimens on a bench, with a short label beside each. At the left end of each row stands a drawn container: a cylindrical vessel with a lid, a tall jar with a foot, a jar with bands and ornament. Between the rows run short paragraphs of text. The whole section is in Currier's language A, by Scribe 1, the scribe of the herbal A pages (Davis, 2020).

THE LAYOUT

The pages are organised as a table. Reading across, a row is a container followed by a series of plant parts; reading down, the containers form a column at the margin. The natural inference, made by everyone since Newbold, is that each row lists the ingredients of something kept in the container, or the preparations made from the plants in that row, and that the paragraphs between rows describe them. The drawings of plant parts are small and schematic but not careless: roots are drawn with their branching, leaves with their lobes, and some can be matched by eye to plants drawn large in the herbal section, though no systematic concordance has been established.

The containers are the most puzzling objects in the book. Some have been compared to apothecary

jars, some to reliquaries, some to the vessels of a distiller. They do not closely resemble surviving pharmacy jars of the period, which were plain earthenware; they are more ornate, with tiers and knobs, and one or two look like pieces of metalwork. They might be drawn from life, from imagination, or from a convention we do not know.

THE LABELS

This is the section where labels matter most, because there are many and they are attached to small, distinct objects. The ZL3b transliteration records 202 words labelling plant parts and 36 labelling containers. The plant-part labels behave like other labels: only one in two hundred begins with **q**, half begin with **o**, and half of the distinct labels also occur as words in the running text. **okol**, **otoldy**, **otoky**, **otaly**, **okaiin** and **oky** each label more than one plant part, and **otaly** and **oky** also label figures in the zodiac.

The container labels are different. There are 36 of them, every one unique, with an average length of seven letters against the five of ordinary words, and only three of the 36 occur anywhere in the running text. Whatever the containers are, their labels are made of the same letters in the same slots, but they are longer and rarer than anything else in the book: **otorchety**, **ofyskydal**, **daramdal**. If any class of word in the manuscript is a set of names, this is the best candidate.

WHAT CAN BE READ OFF THE PAGE

Without reading a word, the section tells us some things. It is a reference list, not a narrative: its unit is the row. It relates plants to containers, and so presumably to preparations, in the manner of a pharmacopoeia or a recipe collection. Its plants are, by the drawings, mostly roots, which is where the active substances of most medieval simples were held to reside. And it was written by the same scribe as the herbal, in the same language, which suggests that herbal and pharmacy were conceived together, as they are in the medical compendia of the period, where a *Tractatus de herbis* is followed by an *Antidotarium*.

Zandbergen has observed that the two bifolia in quire 19 and the one in quire 15 belong together, and that on the style of the containers the quire-19 leaves should come first (Zandbergen, 1996–2026). The section is therefore one more case of a planned sequence broken by the binder.



f. 88v–f. 89r: rows of plant parts, each with a label, and a column of containers at the left of each page. Beinecke Rare Book & Manuscript Library, Yale University.



f. 100r. Beinecke Rare Book & Manuscript Library, Yale University.



f. 99r. The leaves of quire 19, which on the style of the containers precede those of quire 15. Beinecke Rare Book & Manuscript Library, Yale University.

15

Starred Paragraphs

The last quire is the plainest and, by word count, the largest. Twenty-three pages, folios 103 to 116, carry nothing but text: paragraphs of six to twelve lines, each paragraph marked in the left margin by a small star, most of the stars with a tail, some painted red, some with a coloured centre (Zandbergen, 1996–2026). There are 292 such paragraphs in the ZL3b count, counting folio 58 in quire 8, which has the same format and the same stars and is the only other page of its kind. The quire holds 11,602 words, a third of the running text of the whole book, and every page is by Scribe 3 in Currier's B.

A LIST OF SOMETHING

The format is that of a list of entries, and the stars have been read as bullet points, as the mark of a recipe, and as literal stars: one proposal, going back to the 1920s, is that the section is a calendar or almanac with one entry per day or per star, and the association of the zodiac figures with stars gives the idea some colour. The number of paragraphs does not help: 292 is not 365, nor a multiple of 30, nor the number of stars in any catalogue, and two leaves are missing from the centre of the quire, so the original count is unknown.

What the text does say is that these entries are shorter and more uniform than the paragraphs of the herbal section, and that they are in a B that is distinctly its own. Against the bathing section, which is also B and also a single scribe, the recipe pages favour **aiin**, **ar**, **al**, **okeey**, **otaiin**, **chol** and **lkeey**, and use **qokain**, **qol** and **shedy** much less. Over a quarter of all words in the two sections lie in vocabulary

significantly skewed between them. Since the scribe changes as well as the subject, the two causes cannot be separated here as they could for the herbal; but the direction is the same as everywhere else in the book. Different pictures, or no pictures, go with different words.

THE LAST PAGE

Folio 116v, the back of the last leaf, is not part of the sequence. It carries a few lines in a mixture of scripts: words in the Latin alphabet, in a hand that is not that of the main text and may be sixteenth-century, interspersed with words in Voynichese, with small drawings in the margin of an animal and a figure. The Latin-alphabet words have been read in many ways; the readings agree that some of them look German (the sequence usually read as *so nim gaf mich* has been taken as a German phrase) and that others, like the much-quoted *michiton oladabas*, look like nothing in particular. The lines have been interpreted as a charm, a recipe, a pen trial and a later owner's attempt to write in the script (Zandbergen, 1996–2026). They are the only place in the book where the Latin alphabet and the Voynich script share a line, and they have not helped: whoever wrote them was not providing a key.

Two other marginal notes in Latin letters exist, on folios 17r and 66r, both short and both worn; and on folio 49v a column of letters or numerals runs down the margin beside the lines of text, in a hand that may be the same as the month names (Zandbergen, 1996–2026). None has been read with confidence.

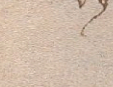
WHAT THE SECTION OFFERS

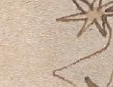
For a decipherer the recipe section is the richest material in the book: a large body of text in a uniform format by one hand, with no pictures to complicate the layout, and with a visible unit, the paragraph, that presumably corresponds to a unit of content. If the entries are recipes, each should name ingredients that also appear in the herbal and pharmaceutical sections, and a decipherment that got the herbal right would be checkable here. If they are almanac entries, the structure should repeat with the calendar. Either way the section is where any claimed reading of the manuscript would have to prove itself at scale.

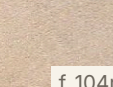




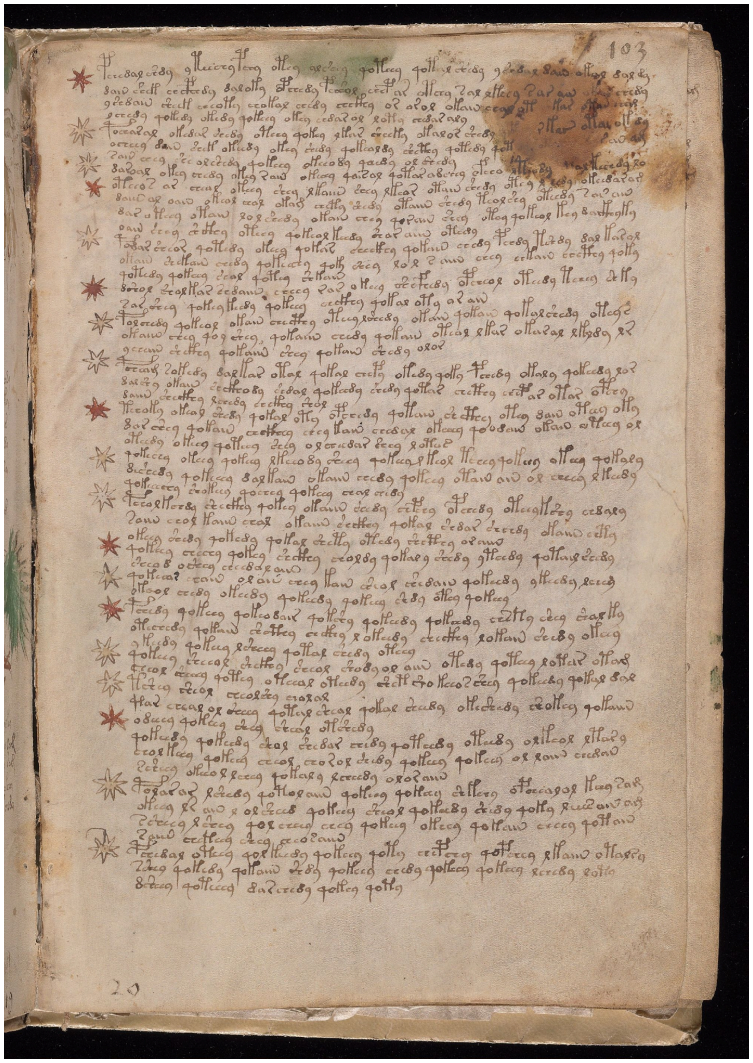




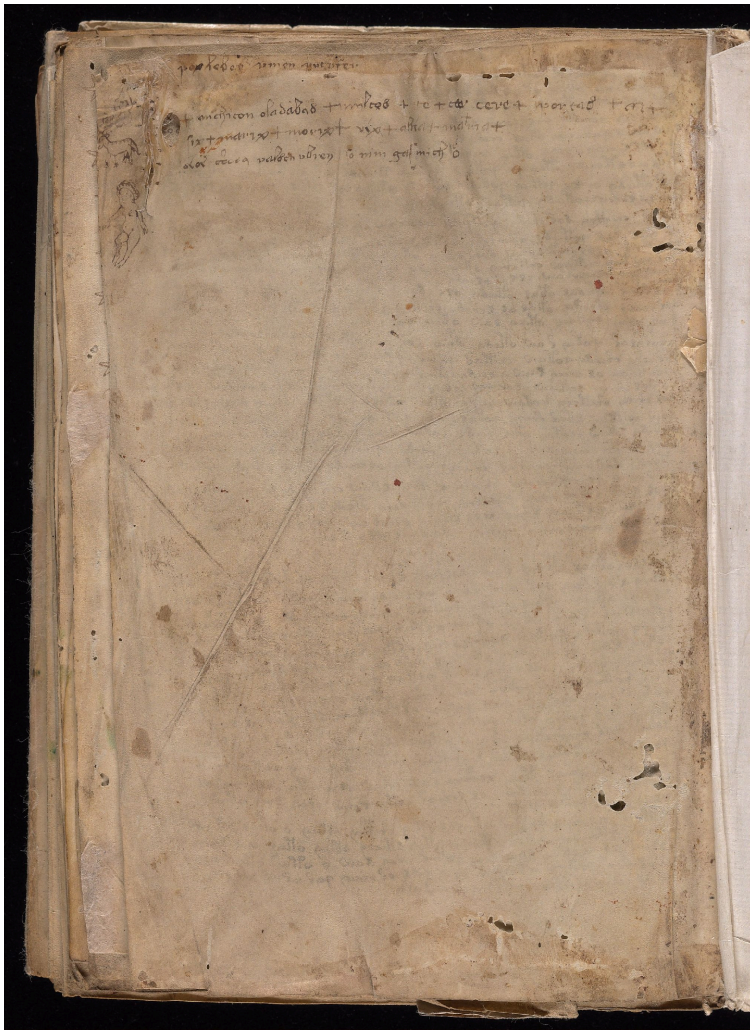




f. 104r. Starred paragraphs: the format of the last quire. Beinecke Rare Book & Manuscript Library, Yale University.



f. 103r, the first page of the quire; its stars, like those of its partner f. 116, lack tails. Beinecke Rare Book & Manuscript Library, Yale University.



f. 116v, the last page: lines mixing Latin letters and Voynichese, with marginal drawings. Beinecke Rare Book & Manuscript Library, Yale University.

Part IV

Explanations

16

What It Could Be

Every explanation of the manuscript belongs to one of four families. The text is a natural language written in an invented script. It is a natural language disguised by a cipher. It is a constructed language or artificial system of some kind. Or it is meaningless, produced to look like writing. Within each family the proposals differ in detail, and a few straddle the lines, but the four are enough to organise the evidence, and this chapter takes them in turn, holding each against what Parts I and II established. The aim is not to pick a winner. It is to see which survive.

NATURAL LANGUAGE, UNKNOWN SCRIPT

The idea is simple: somebody wrote a real language, perhaps their own, in an alphabet of their own devising, and the text is as readable as any language in any script once the script is understood. This is how Linear B and the Maya glyphs were eventually read. It is the explanation most readers assume by default, and it has been given many specific forms: Latin, Hebrew, a Romance dialect, a Turkic language, Nahuatl, a lost language of the Alps.

Three findings bear on it. First, the character-level entropy. A language written out plainly in an alphabet, with one symbol per sound, has the conditional entropy of that language, which for every European language measured lies between about 3 and 3.5 bits. Voynichese has 2.1 to 2.3. No known language written plainly is that predictable; Bowern and Lindemann's survey of 250 found none (Bowern and Lindemann, 2021). So the script, if it is a script, is not a plain alphabet. It could be a syllabary, which writes consonant-plus-vowel as one symbol and so

has fewer choices at each step; or an abjad that drops vowels, though abjads normally raise the entropy, not lower it; or an alphabet with positional variants, in which the same sound is written differently at the start, middle and end of a word, as Arabic letters are. Positional variants would produce exactly the slot structure of chapter 6 (Bowern and Lindemann, 2021).

Second, the word lengths. Natural languages have many short function words; Voynichese has few one- and two-letter words and a tight peak at five. A syllabary or a positional alphabet could produce this if most syllables or words were written with the same number of marks.

Third, the repetition. **daiin daiin daiin** and its cousins are hard to find in natural prose. Some languages reduplicate; some texts, like litanies, repeat. The rate here, ten to twenty times that of the reference texts, would need a genre as well as a language.

The family survives, in its syllabary and positional-variant forms, with the burden of explaining why, in a century, no one has found the language. Our own substitution experiments in chapter 17 show that no simple mapping onto Latin, Italian, German or English works, with or without vowels, but they do not test the languages no one has tried, and they cannot.

CIPHER

The second idea is that the text is a known language, almost certainly Latin or a vernacular of the region, transformed by a procedure meant to hide it. This was the first explanation anyone offered—Baresch, Kircher's correspondents, and the Yale catalogue all call the book a cipher—and the professional cryptanalysts of the twentieth century, Friedman, Tiltman, Currier, all approached it as one, and all failed.

Ciphers of the early fifteenth century are well documented. Simple substitution was the rule; nomenclators, which replace common words and names by symbols, were coming in; transposition was rare; polyalphabetic ciphers were still half a century away. The entropy excludes simple substitution outright. It also makes polyalphabetic substitution unlikely, since such ciphers flatten the statistics rather than sharpening them, and the index of coincidence shows nothing of the kind. A nomenclator alone would leave the ordinary text readable.

What the statistics are compatible with is a verbose cipher: a scheme in which each letter of the plaintext is replaced not by one symbol but by a short group drawn from a table, so that several Voynichese glyphs correspond to one plaintext letter. Such a scheme lowers the entropy, because the glyphs within a group predict one another; it produces words of nearly uniform length, because the groups are of nearly uniform length; and it produces the slot structure, because each group has a shape. Michael Greshko's Naibbe cipher of 2025 is a worked example: using tables that could be chosen with a deck of playing cards, it turns Latin or Italian into text with most of the statistical features of Voynich A, including the slot grammar, and the author is careful to say that it reproduces the features of B less well and that it is a demonstration of possibility, not a decipherment (Greshko, 2025).

A verbose cipher has costs. It inflates the text—a page of Latin becomes two or three pages of

symbols—and the Voynich manuscript, if it is a cipher of this kind, is a short book. It must also account for the line effects, which an enciphering clerk working letter by letter would not produce unless the plaintext had been laid out line by line first. And it must account for the fact that the sections differ in vocabulary in the way real books do, which it does naturally if the plaintext does. The family survives, in this form, with the burden of explaining why the cipher has not been reversed given the quantity of text: thirty-seven thousand words is a great deal of ciphertext for a fifteenth-century system.

CONSTRUCTED SYSTEM

The third family holds that the text is neither a natural language nor a cipher of one but something made: a philosophical language of the kind that seventeenth-century thinkers like Wilkins and Dalgarno designed, a shorthand, a system of glossolalia written down, a notation for something other than speech. Such systems are usually regular to a fault, and the regularity of Voynichese is what makes the idea attractive.

Against it is the date. The great age of invented languages begins two centuries after the manuscript, and no fifteenth-century system of the kind is known, though the *ars combinatoria* of Ramon Llull and the abbreviation systems of the law courts show that the century had the ingredients. Against it too is the vocabulary, which behaves like the vocabulary of a book about things rather than of a category system: words cluster by subject and recur across the pages about one subject. The family is not excluded, but nothing specific supports it, and every specific proposal within it has been a reading from the outside in.

MEANINGLESS

The last possibility is that the text means nothing: that it was made to look like writing, for money, for prestige, or for reasons we would not now understand, by a procedure that produced plausible symbols without encoding anything. Gordon Rugg argued in 2004 that a table of syllables and a card with holes cut in it, the Cardan grille, could generate Voynich-like text quickly (Rugg, 2004); Andreas Schinner found in 2007 that the text has statistical features of a stochastic process (Schinner, 2007); Timm and Schinner proposed in 2020 the autocopying procedure described in chapter 9, in which the scribe writes each new word by copying a word already on the page and altering it, and showed that it reproduces the near-repeat rate, the slot structure and much of the line behaviour (Timm and Schinner, 2020).

This family has one great strength. It is the only one that explains the repetition without needing a genre, and it explains the slot grammar as a by-product of copying. It also fits the fifteenth century: there was a market at courts for books of secrets, and a book that looked like one would have sold.

Against it stand three findings. The first is the long-range order documented by Montemurro and Zanette and confirmed, with the scribe held constant, by our contrasts: the words are distributed across the book as the words of a book about sev-

eral things are (Montemurro and Zanette, 2013). A generator would have to be steered section by section to produce this, which is possible but is a further design decision. The second is the labels: words beside drawings that are drawn from the general vocabulary, are rarely repeated, lack the line-initial marks, and are longest and rarest on the containers. A generator producing labels would have to know that labels are different. The third is the distance profile of the near-repeats. If words were copied from nearby, the rate of near-twins should depend strongly on order. It depends on it a little: among the ten words before any given word, 27 per cent contain a twin in the written order and 24 per cent when the words of the page are shuffled (chapter 17). Most of the near-twin rate is a property of the vocabulary—many words differ from one another by a letter—and survives any reordering. Copying contributes the small remainder at best.

None of these is decisive. The hoax hypothesis can be elaborated to meet each. But each elaboration adds a rule, and the procedure that results—generate words by slots, vary the vocabulary by section, treat labels differently, mark line starts and ends, do it consistently across five scribes—is as complicated as a cipher and harder to teach.

WHERE THIS LEAVES THE QUESTION

Two families are in reasonable health: a natural language in a non-alphabetic or positionally variant script, and a verbose cipher of a known language. One is alive but unsupported: a constructed system. One is alive but strained: meaninglessness. The two healthy families share a feature, which is the most useful thing the evidence has to say: in both, the symbols of the script are not letters in the ordinary sense. Several symbols together correspond to one unit of whatever is underneath, and the slots of chapter 6 are the seams between those units.

That is where chapter 17 begins.

17

Attempting a Reading

This chapter reports what happened when we tried to read the text. We set out the experiments in the order we ran them, with their results, including the negative ones, because a negative result with its method attached is worth more than a positive result without. Everything here can be rerun from the scripts in Appendix A in about a minute.

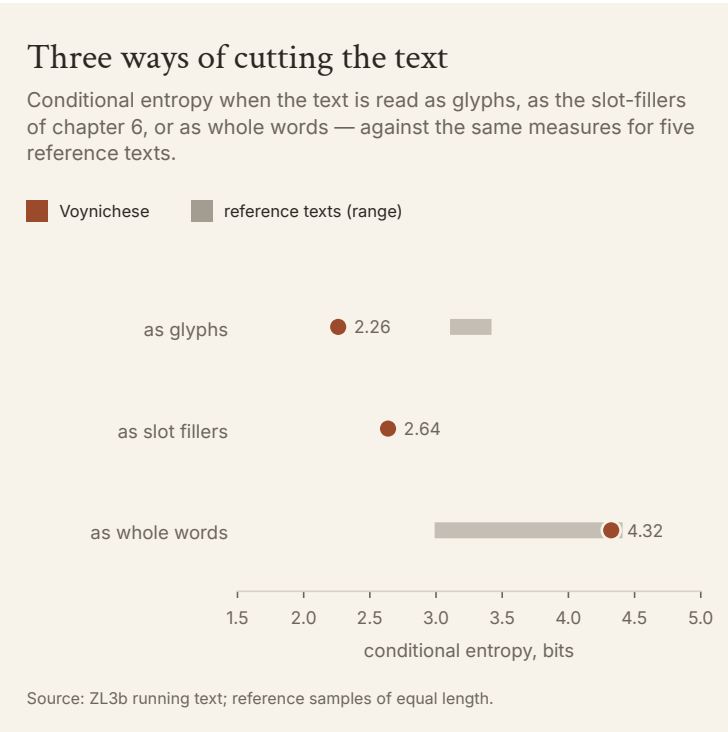


Figure 17.2 Entropy at three levels of cutting.

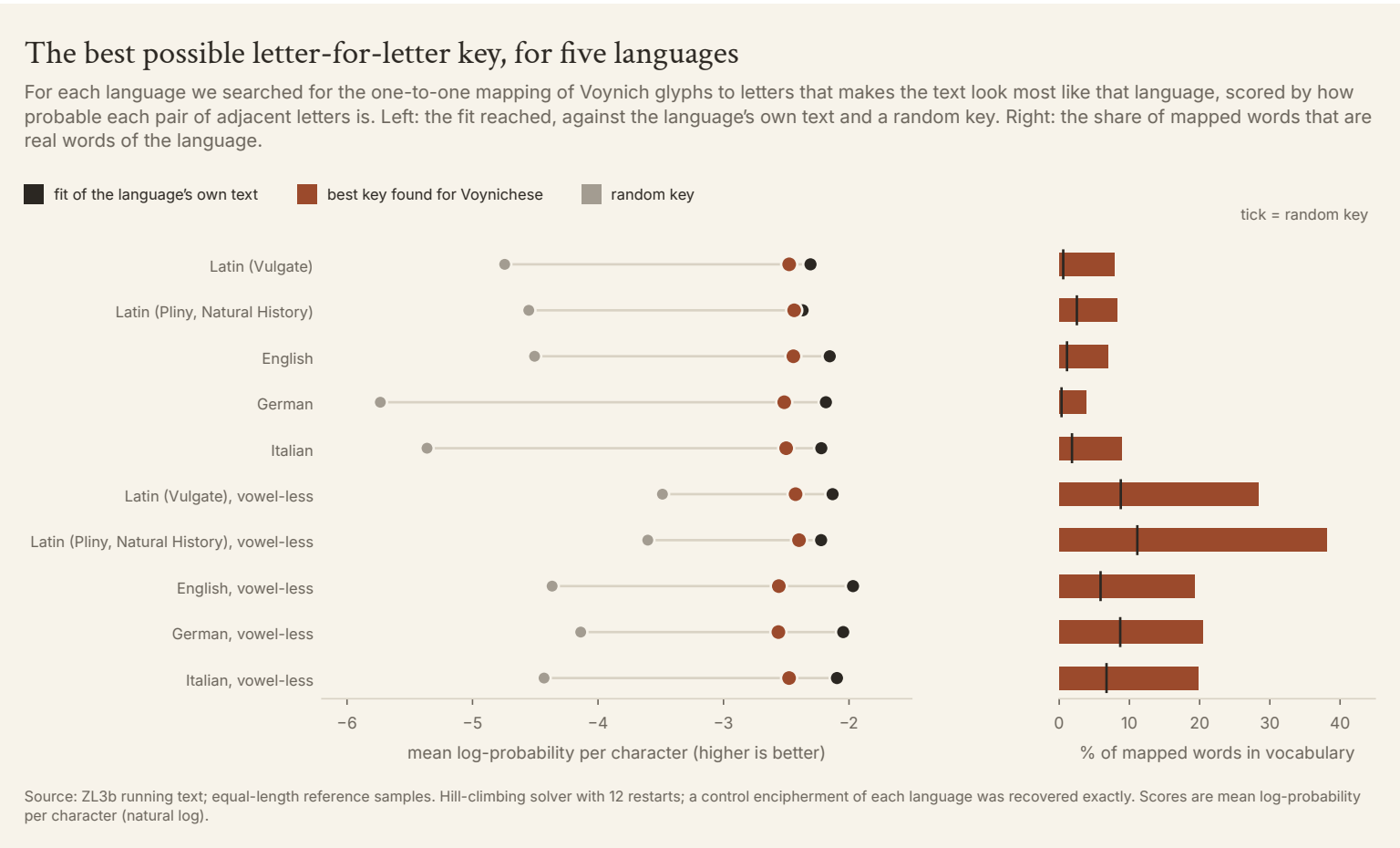
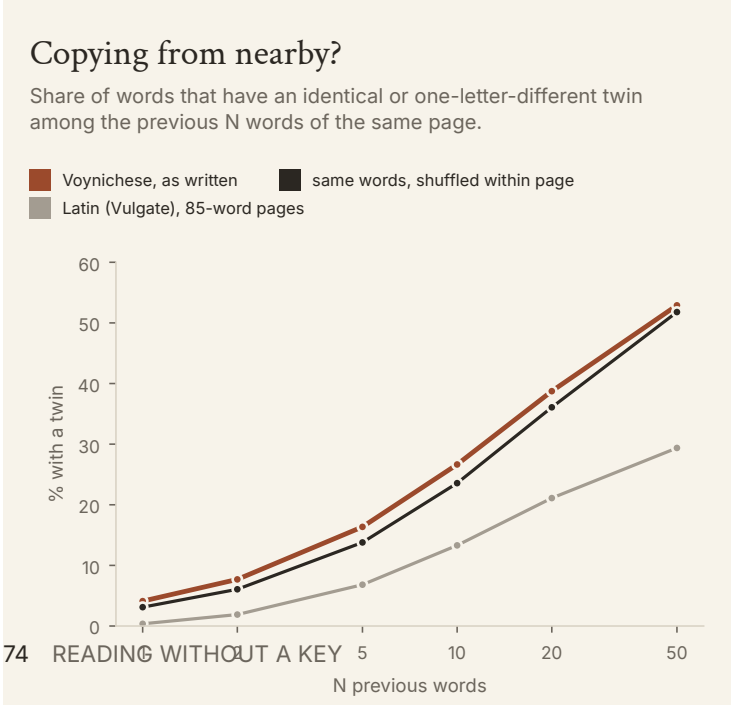


Figure 17.1 The best one-to-one keys and what they produce.



EXPERIMENT 1: THE BEST POSSIBLE KEY

The simplest cipher hypothesis is a one-to-one substitution: each Voynich glyph stands for one letter of some language. Chapter 9 argued from the entropy that this cannot be right. We tested it directly anyway, because the argument from entropy is indirect and because the test produces something the entropy does not: the best key there is, and what the

text looks like under it.

The method is standard. For a candidate language, build a table of how often each letter follows each other letter in a sample of that language. Then search for the assignment of Voynich glyphs to letters that makes the Voynich text score highest under that table: start from a random assignment, swap two letters, keep the swap if the score improves, repeat several thousand times, restart from scratch a dozen times, keep the best. The glyphs were the EVA units with the benches and gallows-benches as single symbols, the twenty-two to twenty-five commonest, mapped one-to-one onto the letters of the language. The word space was fixed to the word space.

The search works. As a control we took each reference text, enciphered it with a random substitution, and gave the result to the same search; in every case the search recovered a key whose score was identical to the plaintext's own, that is, it found the plaintext. Figure 17.1 shows what it found for Voynichese. For every language the best key scores far above a random key and below the language's own text. The gap to the language's own text is smallest for Pliny's Latin (0.07 nats per character) and largest for German (0.33). A gap of 0.07 nats per character sounds small; over 190,000 characters it is a likelihood ratio of $e^{12{,}000}$, which is to say that the Voynich text, under the best Latin key, is not Latin.

What the best keys produce is shown in the right-hand panel of the figure. Under the best Latin key, 8 per cent of the mapped words are words that occur in the Vulgate; under a random key, under 1 per cent. Under the best English key, 7 per cent; German, 4; Italian, 9. The first line of the manuscript under the best Vulgate key reads *xuaed etus um ubullo pis pime hmrd e tim pisce dime gum*. The hits are short words that any assignment of common letters will produce by chance, *um*, *e*, *tim*, and they do not join up.

EXPERIMENT 2: WITHOUT VOWELS

Medieval scribes abbreviated heavily, and one family of proposals holds that the text is a language with its vowels dropped, as Hebrew and Arabic drop them. We repeated the search against vowel-less versions of the five reference texts. The fit is worse, not better: the gaps widen to between 0.18 and 0.59 nats per character. The dictionary hit rates rise, to 19–38 per cent, but so do the random-key rates, to 6–11 per cent, because a vowel-less vocabulary is full of two- and three-consonant strings that almost any mapping will produce. The output is *?bdtg thbm bs bvbllq frm frst*. We do not regard this as a test of the abjad hypothesis in general, only of its simplest form.

EXPERIMENT 3: THREE WAYS OF CUTTING

Chapter 16 concluded that in the surviving explanations the glyphs are not letters, and that several glyphs make up one underlying unit. If so, the right measure of the text's predictability is not glyph-by-glyph but unit-by-unit, and the question is whether, cut at the right joints, the text looks like language.

We cut it three ways (figure 17.2). As glyphs, the conditional entropy is 2.26 bits, far below the 3.1–3.4 of the references. As slot-fillers—each word divided by the template of chapter 6 into its opening, gallows, middle, and ending, and each filler treated as one symbol—it is 2.64 bits, with 43 distinct fillers.

That is closer but still below the natural range, which suggests either that the slots are not the units, or that the units are larger than single slots. As whole words it is 4.32 bits, and here the picture changes: the reference texts, cut into words, range from 3.0 to 4.4. At the level of the word, Voynichese is exactly as predictable as the Vulgate, the King James Bible or the Luther Bible.

That is a real result, and it should be stated carefully. It means that whatever process produced the sequence of words produced one with the same amount of order from word to word as a text in a language has. It does not mean the words are words of a language; a nomenclator, a syllabary with word-sized groups, or a well-steered generator could do the same. But it locates the level at which the text behaves like language: not the glyph, not the slot, but the word.

EXPERIMENT 4: IS ANYONE COPYING?

Timm and Schinner's hypothesis makes a prediction about distance: if a scribe wrote each word by copying and modifying a word already on the page, words should have their near-twins nearby. We measured, for every word on every page with thirty or more words, whether an identical or one-letter-different word occurred among the previous one, two, five, ten, twenty or fifty words, and compared the result with the same pages with their words shuffled (figure 17.3).

The written order has more near-twins than the shuffled order at every distance, but not many more: 4 per cent against 3 at a distance of one word, 16 against 14 at five, 27 against 24 at ten, 53 against 52 at fifty. The Latin control, cut into 85-word pages, is far lower at every distance (0.4, 7, 13, 29 per cent). The near-twin rate is mostly a property of the vocabulary, present in any order; the written order adds one to three percentage points. Something local is going on—the adjacent repeats of chapter 9 are real—but copying from the surrounding lines is not the main engine of the text's repetitiveness, and a procedure that was would show a steeper curve.

EXPERIMENT 5: CRIBS

A crib is a piece of known plaintext, and cribs are how most historical ciphers were actually broken. The manuscript appears to offer several. The month names on the zodiac pages are not cribs: they are in a later hand, in Latin letters, and attached to the pictures, not the text. The labels are the real candidates. If the thirty labels around the Pisces wheel were the names of the thirty degrees of Pisces, or the names of thirty stars, in some language, one could look for a mapping that turned them into a known list.

We tried the obvious structural test. Names of the degrees of one sign would share nothing in particular with the degrees of another; names of stars would be unique. The 281 distinct zodiac labels include 37 that recur on other signs, four of them on four signs apiece, and they are built from the common vocabulary of the running text. Within a sign the labels do lean toward a shared opening: 71 per cent of the labels on the second Aries wheel (f. 71r) begin with **ot**, 52 per cent of the Gemini labels with **ok**, and the **ot**-heavy signs are the early ones, Pisces, Aries and

Taurus, while the later signs favour **ok**. But **ot** and **ok** are the two commonest openings of label words everywhere in the book, and no sign has an opening or ending of its own. The labels look, in short, like ordinary words of the text standing beside figures, with a drift in their spelling from one end of the zodiac to the other, not like a list of names. The pharmaceutical container labels look more like names: unique, long, and absent from the running text; but there are thirty-six of them, beside containers that cannot be identified, which is not enough to fix a key.

EXPERIMENT 6: OTHER PEOPLE'S KEYS

We did not repeat the published decipherments, because each has already been examined and found wanting on the same grounds, which are worth stating once. Newbold's 1921 reading depended on marks that are ink cracks. The readings of the 2010s and 2020s that have reached the press—Stephen Bax's partial identifications of plant and star names, Nicholas Gibbs's 2017 claim of abbreviated Latin, Gerard Cheshire's 2019 "proto-Romance," and several since—share a method: assign values to a few glyphs so that a few labels become words in a chosen language, then extend the values to a line of text and translate the result freely. Each produces readings that do not transfer: the key that makes one label a plant name makes the next label nothing, and the running text under the same key is not language in any sense that a second reader can verify (Bowern and Lindemann, 2021; Davis, 2017). Hauer and Kondrak's 2016 computational study, which treated the text as an anagrammed abjad and found Hebrew the best-scoring of many languages, is a different kind of work and its authors were careful about it; the fit was weak, the assumption of anagramming was unmotivated, and the decoded sample did not withstand a reading by Hebrew speakers (Hauer and Kondrak, 2016).

WHAT WE HAVE, AND WHAT WE DO NOT

We have not read the text, and we do not think anyone has. What the experiments add to the picture is this. No one-to-one key to a European language exists, with or without vowels; the best such keys produce no language. The text is as orderly as language at the level of the word and far more orderly than language at the level of the glyph, which places the seam between symbol and meaning at the word or just below it. The repetition is in the vocabulary, not the copying. And the labels, which ought to be the cribs, are made of the same words as everything else.

If a reading is to come, we think it will come from the labels and the recipes together, with the sheets in their original order, and with a candidate system—a verbose cipher or a syllabary—whose units are word-sized. Chapter 19 says what such a reading would have to show.

18

A Defensible Reading

A Reading, in the sense this book has used the word, is an account of what is on the page that can be defended line by line. This chapter gives ours. It is organised from the most secure statements to the least, and each paragraph says where its support comes from. Nothing in it is a translation.

WHAT THE BOOK IS

It is a single project. The parchment was prepared in one batch in the first third of the fifteenth century; the quires were numbered by one fifteenth-century hand; the five scribes share one script, one set of drawing conventions and one page design; the two languages are distributed across scribes and sections as if assigned. A book made over many years by unrelated people, or assembled from parts, would not look like this. (Chapters 2, 4, 8.)

It was made by a group. Five hands is Davis's count, and the alternation of two of them sheet by sheet in the herbal shows that at least two worked on the same section at the same time, each to a plan. The sheets being out of order shows that the binder was not one of them. (Chapters 4, 8.)

It is organised by subject, as a compendium. Plants with text; stars and the zodiac with labels; bathers with text; a cosmological diagram; a pharmacy with labelled parts and containers; a long list of starred entries. These are the parts of a medical and natural-philosophical compendium of the kind that fifteenth-century physicians and apothecaries owned, in which a herbal, a calendar, a treatise on regimen and a book of recipes were bound together. No surviving compendium has these pictures, but ev-

ery section has a genre to belong to. (Chapters 10–15.)

Its sequence is lost. The present order of the sheets is not the maker's, and the colouring was added, at least in part, after the order was lost. Any reading that depends on the order of pages is not yet possible. (Chapter 4.)

WHAT THE TEXT IS

It is a consistent system. Thirty-seven thousand words follow the same rules of word formation, line formation and paragraph formation, across five scribes and two languages. The rules are not those of any known alphabet, and they were not improvised. (Chapters 6, 7.)

Its units are not letters. The glyphs are too predictable from one another to be the letters of a language written out plainly; they are predictable in exactly the way the parts of a composite symbol are. The level at which the text has the orderliness of language is the word, or just below it. (Chapters 9, 17.)

Its two languages are two settings of one system. A and B differ in a handful of affixes and in the frequency of common words, not in their structure; the same scribe writes both subjects in the same language differently, and two scribes write the same subject in the same language alike. (Chapter 8.)

It is about things, or made to look as if it were. The vocabulary of each section is its own, in proportion to the way real books on several subjects divide their vocabulary, and this holds with the scribe held constant. This is the strongest evidence that the text carries content, and it is evidence, not proof. (Chapters 8, 9.)

It is not a one-to-one substitution of Latin, Italian, German or English, with or without vowels. (Chapter 17.)

It is not well explained as meaningless text generated by copying. The repetition that the copying hypothesis was built to explain is a property of the vocabulary, not of the sequence; the section-by-section vocabulary and the behaviour of labels require more of a generator than copying provides. (Chapters 16, 17.)

WHAT THE LABELS ARE

They are words of the same system, without its linemarks. They are mostly unique, mostly **o**-initial, and mostly ordinary vocabulary. Thirty-seven zodiac labels recur on more than one sign, which argues against their being proper names of degrees or stars. The container labels in the pharmacy are the one class that behaves like names. (Chapters 7, 11, 14.)

WHAT THE PICTURES ARE

They are competent, provincial, and conventional in kind without being copies of any known exemplar. The plants belong to the illustrated-herbal tradition; the zodiac emblems to the German tradition of about 1400–1500, with a crossbow Sagittarius; the bathers to the balneological tradition, with a resemblance to the manuscripts of the *De balneis Puteolanis*; the fortification in the rosettes to northern Italy and the Alps. Taken together they place the book's makers in the German-speaking or northern Italian region in the generation after 1404, and no more precisely

than that. (Chapters 10–13.)

WHAT WE THINK, AND HOW FIRMLY

Of the explanations that survive the evidence, we find the verbose-cipher family the most probable, by a modest margin, for three reasons: it explains the entropy, the word lengths and the slot structure as consequences of a single mechanism; it fits the century, in which substitution tables and nomenclators were the ordinary tools of secrecy; and a worked example now exists that reproduces most of the statistics of language A from Latin or Italian plaintext (Greshko, 2025). We hold this at the level of a working hypothesis, not a finding. The natural-language-in-a-syllabary family is close behind and cannot be separated from it by the tests we know how to run.

If the first hypothesis is right, then there is a plaintext, probably in Latin or a Romance or German vernacular, probably a medical compendium, and the key is a table that maps letters to glyph groups of the kind the slot grammar describes. That table has resisted a century of attack, which means either that it is more complicated than the slot grammar suggests—with homophones, nulls and context rules—or that something about the plaintext, such as heavy abbreviation, makes the ciphertext short of the regularities that cryptanalysis needs. Both are possible, and both would be helped by the one thing nobody has yet had: the sheets in their original order.

WHAT WE DO NOT KNOW

We do not know the language, the author, the place, the purpose, or the meaning of a single word. We do not know whether the text is meaningful. We do not know the original order of the leaves, the original number of leaves, or when the book was first bound. We do not know what the women in the water are doing, what the rosettes represent, or which plants the plants are. We do not know what the containers contain. We have set out what would change this in the next chapter.

19

Judging the Next Solution

SOMEONE will announce a decipherment of the Voynich Manuscript within a year of this book appearing; someone does most years. This chapter is for the reader who wants to know whether to believe it. The tests below are not ours. They are the tests that every successful decipherment of an unknown script has passed, from Champollion's hieroglyphs to Ventris's Linear B, and that every failed reading of this manuscript has not. A claimed solution that meets them deserves attention. One that does not can be set aside without discourtesy, however distinguished its author.

THE TESTS

1. It is stated in advance and applied mechanically. The key—the table from glyphs to letters, or words to meanings, or whatever it is—must be written down completely before any passage is read, and reading must consist of applying it. A key that is adjusted as one goes, or that leaves the reader free to choose between alternatives at each word, is not a key. Newbold's reading failed here; so did every "partial decoding" of the past decade.

2. It is applied by someone else. A second person, given the key and a page the proposer has not read, must produce the same reading. If the proposer is the only one who can make it work, it is not working.

3. It produces language. The output must be a text in an identifiable language that native readers or competent scholars of that language accept as grammatical and meaningful, without the proposer's gloss. "Proto-Romance" that no Romance philologist recognises, or Latin that is a string of abbrevi-

ations each expanded as convenient, does not count.

4. It reads the running text, not only labels. Nine tenths of the manuscript is paragraph text. A reading of a dozen labels is a hypothesis about a dozen words. The test is a page of the recipe section, or a herbal paragraph, read in full.

5. It is consistent with the pictures. The herbal paragraphs should discuss plants; the recipe entries should contain ingredients that also appear in the pharmacy; the bathing text should be about what the bathing pictures show, whatever that is. A reading in which the page about a plant turns out to be about something unrelated has a problem.

6. It explains the statistics. The key, applied in reverse to a real text in the claimed language, should produce text with the properties measured in Part II: the slot structure, the entropy, the line effects, the two languages. If the proposed system could not have produced the manuscript, it did not.

7. It is consistent with the object. A reading that requires the book to be sixteenth-century, or Mexican, or the work of one person, has to overcome the radiocarbon date, the pigments, the five hands and the quire marks. These are not interpretations and they do not move.

8. It accounts for its own novelty. If the language is one that a century of experts failed to see, the reading should explain what they missed and why it is only visible now. "Nobody tried X" is sometimes true. It is rarely true of the obvious languages.

WHAT WOULD MOVE THE QUESTION

Short of a decipherment, several pieces of work would change what can be said, and most of them are within reach.

The order of the sheets. Layfield and Davis have proposed to use the water stain to recover the original sequence (Layfield and Davis, 2026); the sewing holes, the paint transfers and the picture across folios 78v and 81r are further constraints. A reconstruction tested against the text's own vocabulary continuity would be the first fully independent check on whether the text has content.

A single transliteration with uncertainty. The two main transliterations agree on three words in four. A consolidated version that records the disagreements as alternatives, rather than choosing, would let every statistic in this book be given with an error bar.

The glyph inventory settled. Davis's digital palaeography has begun this (Davis, 2020). A published, image-linked inventory of distinct forms, with their variants, would remove the largest known source of disagreement between analyses.

A verbose-cipher search. Greshko's cipher shows that a table of the right kind can produce the statistics (Greshko, 2025). The reverse problem—searching the space of such tables for one that turns the manuscript into language—has not been attempted at scale. It is a large computation but a definable one.

The containers. Thirty-six unique long labels beside thirty-six drawn vessels are the best cribs in the book. Identifying the vessels, or finding their models in the material culture of the period, is work for

a historian of pharmacy that has not been done.

The comparanda. The bathers, the rosettes and the zodiac wheels have each been compared with a few manuscripts. Systematic comparison against the digitised corpus of fifteenth-century medical and astrological manuscripts, now large, might find the models the drawings were made from, and a model with text would be a crib of a different kind.

None of this requires reading the manuscript. All of it would make reading it more likely, and all of it is the kind of patient, checkable work that the history of the subject has had too little of. The book has waited six hundred years. It can wait for the order of its pages.

Part V

The Attack

20

A Framework for Testing

The first four parts of this book described the manuscript and set out what the evidence supports. This part reports a second phase of work, undertaken after the first was drafted, with a different aim: not to describe the text but to attack it, systematically, with every instrument that seemed worth borrowing from a neighbouring field, and to record the results in a way that could not be quietly adjusted afterwards.

The history of the subject makes the second requirement the important one. Nearly every failed solution of the manuscript began as a reasonable idea tested by its author, who then kept testing until something fit. The remedy is the ordinary scientific one, applied with more than ordinary strictness because the material invites self-deception. Before any test in this part was run, its hypothesis, its predictions, the comparison that would count as a null result, and the rule for deciding were written down in a registry. The registry and the test outputs are distributed with the book; Appendix E prints the registry as it stood when the book went to press.

HYPOTHESES AND PREDICTIONS

A hypothesis, for our purposes, is a claim about how the text was produced that implies at least one measurable prediction. "The text is Latin in a cipher" qualifies. "The text is meaningful" does not, until it is attached to something that could be observed if it were true and not if it were false; and several of the proposals in the literature, when pressed, turn out to have no such attachment. Those we list as unfalsifiable and do not score. Each hypothesis in the registry carries its predictions in the form of a

named statistic and an expected direction, and a note on what would count against it.

NULL MODELS

A statistic on its own means nothing; it needs something to be compared with. We use three kinds of comparison throughout, and every result in the chapters that follow carries at least one.

The first is a *surrogate*: the manuscript's own text with its structure destroyed at one level. Shuffle the letters within each word and the words survive but their spelling is lost. Shuffle the words within each line and the line survives but word order is lost. Shuffle words within pages, or across the whole book; shuffle lines across the book; shuffle pages. Each surrogate removes structure at one scale and keeps it at the others, so that a statistic computed on the real text and on each surrogate shows at which scale the effect lives.

The second is a *reference distribution*: the same statistic computed on samples of 103 languages, each the same length as the manuscript's running text. The samples come from the parallel Bible corpus assembled by Christodoulopoulos and Steedman, which has the great advantage for this purpose of being the same text in every language, so that genre is held constant (Christodoulopoulos and Steedman, 2015). For the 84 languages written in alphabets we add a vowel-less version, to stand in for the abjads and for the abbreviation hypotheses. "Unusual" in this part means outside the range of those 187 samples, not merely different from Latin.

The third is a *positive control*: a text produced under the hypothesis being tested, on which the test must succeed before its result on the manuscript is allowed to count. Every solver used in chapter 23 was first given Latin, Italian and German enciphered with keys we chose, and had to recover them. A solver that cannot find a planted key is not evidence of anything when it fails to find one in the manuscript.

DECISION RULES

Rules were fixed per test before running it. For comparisons with the reference distribution, a value outside the central 95 per cent of the 187 samples counts as unlike attested languages. For solvers, a candidate key counts as a hit only if it transfers to pages not used in fitting it, if the share of decoded words that are real words of the claimed language exceeds the share a random key produces by more than the control's margin, and if a scorer not used in the search agrees. Where the same test is run against many languages, the best candidate is reported with its margin over the next, never alone.

HELD-OUT SHEETS

Any key or model fitted to the text is fitted on one set of quires and scored on another, and the split is by quire rather than by page, because pages on one sheet share scribe and subject and would leak. The training set is quires 1, 3, 5, 7, 13, 15 and 20; the rest is held out. Labels are never used to fit a key that is then scored on labels.

WHAT COUNTS AS PROGRESS

Progress toward a reading is counted in words of running text that a proposed key renders as dictionary words of the claimed language on held-out sheets, above the rate a random key achieves. Labels alone do not count, for the reasons given in chapter 7. By this measure, as the following chapters report, the count at the time of writing is zero. The framework is designed so that the first non-zero count, when it comes, will be one that a sceptic can check.

21

One Hundred Languages

The five reference texts of Part II were enough to show that Voynichese is not written like Latin or German. They were not enough to say what it is written like, because five languages are five points, and the world has seven thousand. This chapter widens the comparison to 103 languages in 187 samples, asks on which statistics the manuscript is unlike any of them, which of them it most resembles and at what level, and whether any of them, under a one-to-one substitution, could be the text.

THE PROFILE

Figure 21.2 is the result in one picture. Each row is one statistic computed on the manuscript's running text and on every language sample; the grey band is the range within which 95 per cent of the languages fall. On thirteen of fourteen statistics the manuscript lies inside the band. On one it does not: the conditional entropy of the characters, the measure of how predictable each symbol is from the one before. The lowest value among the 103 languages, plain or vowel-less, is 2.51 bits; the manuscript's is 2.13, or 2.26 with the benches counted as single symbols. Bowern and Lindemann found the same against 250 languages (Bowern and Lindemann, 2021); it is the one robust anomaly at the level of the symbol, and the comparison with vowel-less forms matters, because dropping vowels is the obvious way to lower a text's entropy and it does not go far enough.

The other thirteen statistics tell a more interesting story, because they place the manuscript in different company depending on the level at which one

looks. Take the character-level statistics alone—symbol entropy, word length, the spread of word length, the share of one- and two-letter words—and the nearest languages are Malagasy, Akawaio, Maori, Cebuano and Tagalog: languages with simple open syllables, consonant-vowel-consonant-vowel, written in Latin letters (figure 21.1). Take the word-level statistics alone—distinct words per word, the share of words used once, the Zipf slope, the share of the commonest words, the conditional entropy of the word sequence—and the nearest are Quechua, Swahili, Zulu, Nahuatl and Ashaninka: agglutinative languages, in which a word carries several morphemes and few words recur exactly.

Nearest languages, by level

The ten languages whose statistical profile is closest to Voynichese when only character-level statistics are used (left) and when only word-level statistics are used (right). Distance in z-score units.

character-level		word-level	
Malagasy	1.9	Aguaruna (no vowels)	1.3
Akawaio	2.1	Zulu (no vowels)	1.4
Uma	2.3	Swahili	1.5
Maori	2.3	Xhosa (no vowels)	1.6
Camsa (no vowels)	2.3	Ukranian	1.6
K'iche'	2.3	Ashaninka (no vowels)	1.6
Galela	2.4	Shuar	1.6
Cebuano	2.4	Ashaninka	1.6
Tagalog	2.4	Syriac	1.6
Cakchiquel	2.4	Quichua	1.8

Source: profile screen, 103 languages × plain/vowel-less. ‘novowel’ = vowels removed.

Figure 21.1 Nearest languages by level.

Neither list is a candidate language; the manuscript’s symbol-level behaviour has already excluded any plain alphabet. What the two lists say is structural. At the level of the symbol, Voynichese looks like a script with a small, regular syllable: a syllabary, or a cipher whose units are syllable-sized. At the level of the word, it looks like a language, or a code, in which a “word” is a bundle of several pieces and bundles rarely repeat. Both observations are consistent with the slot structure of chapter 6 and with the verbose-cipher family of chapter 16. They are also the clearest statement we can make of what any successful explanation has to look like: syllable-sized units, assembled several at a time into words.

ONE HUNDRED AND EIGHTY-SEVEN KEYS

The one-to-one substitution search of chapter 17 was run against five languages. We ran it against all 187 samples, with the design tightened: the key is fitted on the training quires only and scored on the held-out quires; a planted key in each language is solved first as a control; the share of decoded held-out words that are dictionary words of the language is compared with the share under random keys.

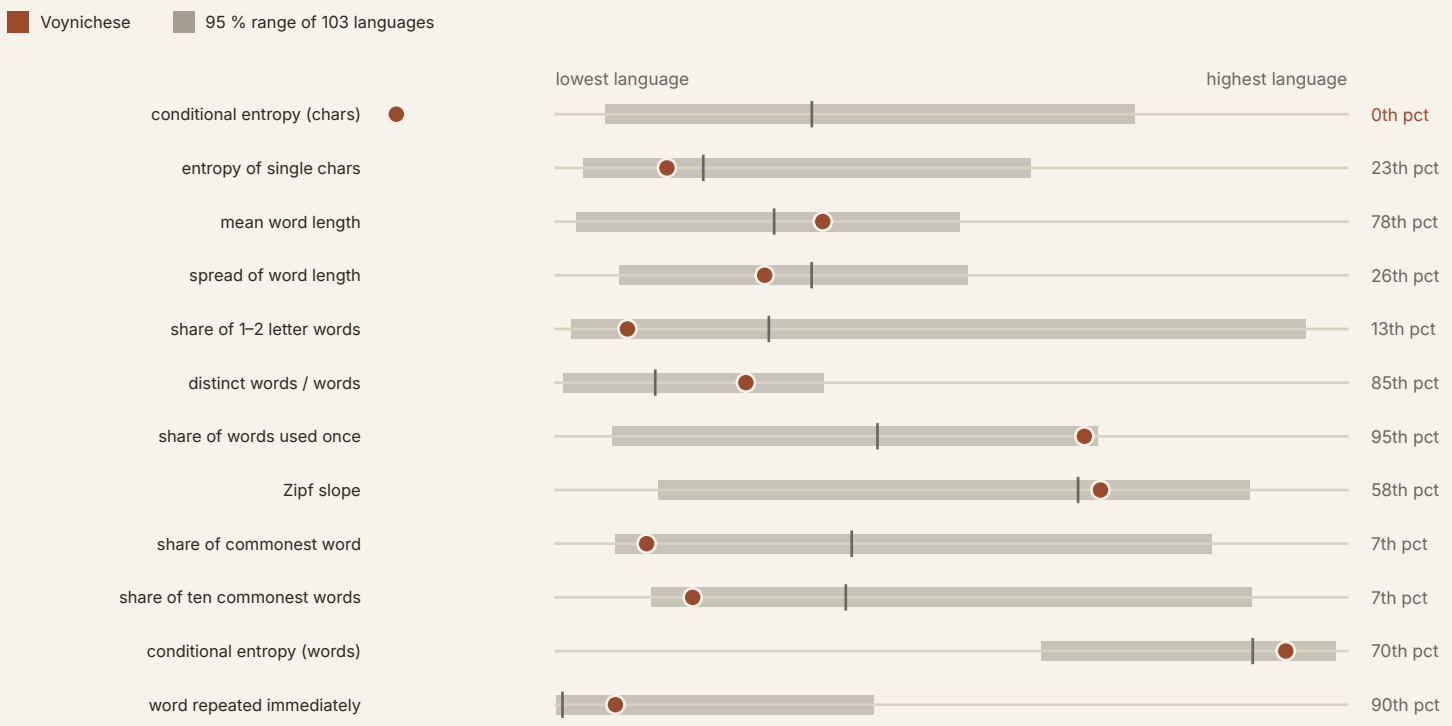
No language passes. The smallest gaps between the manuscript’s best key and the language’s own text are for Arabic (0.10 nats per character), Hebrew (0.11) and Syriac (0.14), with Thai and the Berber language Tachelhit next. The dictionary rates for the three Semitic languages are 7–15 per cent against 2–5

per cent for random keys, and the decoded text is not text. The proximity of the abjads is real and worth recording: a text written without vowels has fewer symbols and tighter sequences, and the manuscript's bigram structure fits such a text better than it fits a full alphabet. It is the same signal that led Hauer and Kondrak to rank Hebrew first in 2016 (Hauer and Kondrak, 2016), and it has the same limitation: "fits an abjad better than an alphabet" is a statement about entropy, not about Hebrew. The vowel-less samples show high dictionary rates under the best keys, 20–50 per cent, but their random-key rates are 5–27 per cent, because a vowel-less vocabulary accepts almost any string of consonants; the margins are no better than the Semitic ones.

The conclusion of chapter 17 therefore stands on a wider base: there is no one-to-one key from the glyphs to the letters of any of 103 languages, in plain or vowel-less form, that produces that language on pages it was not fitted to. Whatever the relation between the glyphs and the underlying text, it is not one symbol to one letter.

Where Voynichese sits among 103 languages

Each row is one statistic. The grey band spans the 2.5th–97.5th percentile of 103 Bible-corpus languages in equal-length samples; the tick is the median; the dot is the manuscript. Only one statistic falls outside the attested range.



Source: ZL3b running text; Christodoulopoulos & Steedman Bible corpus, 34,180-word samples per language. Percentile position shown; values in data/framework/out/profile_screen.csv.

Figure 21.2 The manuscript's statistical profile against 103 languages.

22

Borrowed Instruments

A symbol sequence of unknown meaning is not a problem unique to manuscripts. Molecular biologists read genomes without a dictionary; physicists characterise signals without knowing their source; cryptanalysts, linguists and students of animal communication all have instruments for asking of a sequence what kind of thing produced it. This chapter applies several of those instruments to the manuscript, each with the null models of chapter 20, and reports what they add. Most add a little. Two add something new.

HOW FAR DOES ONE SYMBOL REACH?

The conditional entropy of chapter 9 measures how much the next symbol depends on the current one. Mutual information at a distance generalises the question: how much does a symbol tell you about the symbol d positions later, for d from one to sixty? In genomes this profile distinguishes coding from non-coding regions; in texts it shows how far structure extends (Li and Kaneko, 1992; Montemurro and Zanette, 2013). Figure 22.1 gives the manuscript's profile against eight languages and against its own surrogates.

At a distance of one the manuscript is far above every language: 1.6 bits against 0.6 to 1.3. By a distance of five it has fallen to the languages' level, and by ten it is within a whisker of the surrogate in which the words have been shuffled across the whole book. The surrogate in which words are shuffled only within their lines sits almost exactly on the real text at every distance. In plain terms: the manuscript's order is inside the word and inside the line, and there is little beyond the line that the sym-

bols can detect. This is consistent with chapter 7 and sharpens it: whatever the line is, it is close to the largest unit of which the glyph sequence keeps track.

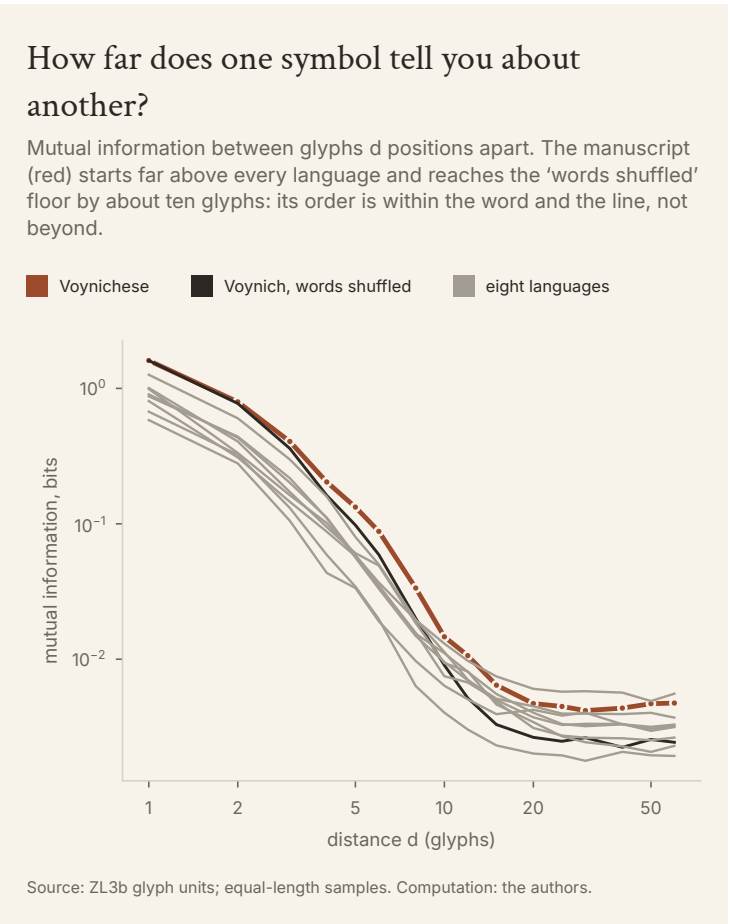


Figure 22.1 Mutual information at a distance.

PHRASES

Chapter 9 reported that the manuscript repeats individual words far more often than languages do. The bioinformatician's next question is about longer repeats: are there passages that recur, as formulae recur in recipes, prayers, or lists? We searched for every sequence of three or more words that occurs more than once, and compared the count with the same search on equal-length samples of nine languages cut into lines of the manuscript's length.

The result (figure 22.2) is the reverse of the word-level finding. The manuscript has 97 repeated three-word sequences and not one repeated four-word sequence. Pliny's Latin, the least formulaic of our references, has 128 and 16; the Bible texts have hundreds and thousands. A text that is formulaic at the level of the word is, at the level of the phrase, less repetitive than any prose we could find. Two explanations fit. One is that the text is a list of distinct items in which phrases are not expected to recur, which is true of some herbals and inventories but not of recipe collections. The other is that a homophonic cipher is at work: when each plaintext letter can be written several ways, two writings of the same phrase almost never coincide, while short words, which have fewer letters to vary, coincide often. The second explanation makes a further prediction, taken up below.

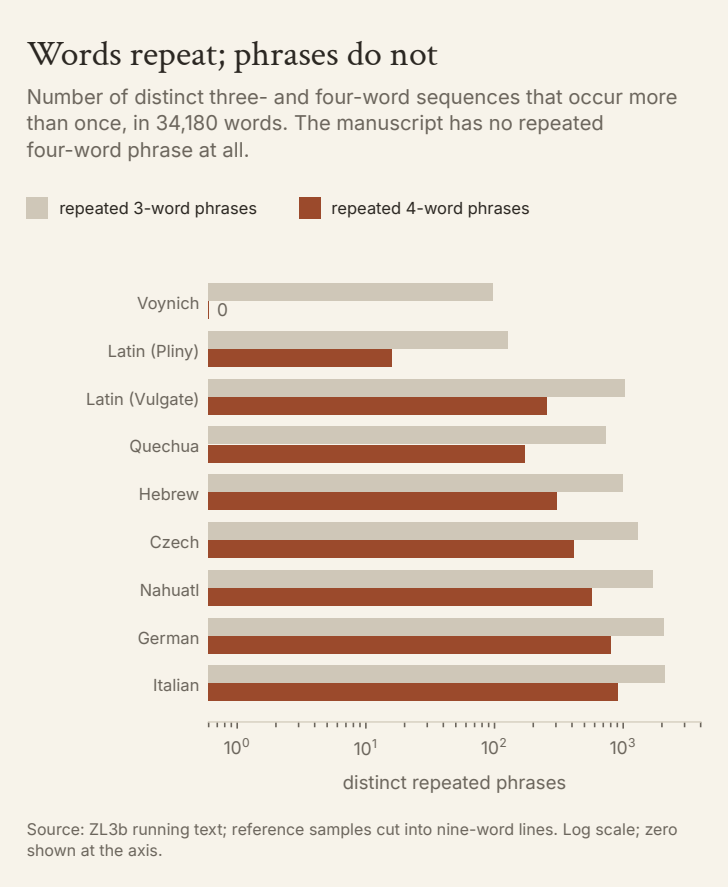


Figure 22.2 Repeated phrases.

COMPRESSION AND MERGES

Two simpler measures agree with the entropy. The text compresses to 31 per cent of its size with a standard algorithm, against a median of 27 per cent for the languages, which is to say it is somewhat *more* compressible than its own words shuffled and in the upper third of the language range: redundant, but not extraordinarily so once the word level is included. And when the commonest adjacent pairs of symbols are merged one at a time, as a tokeniser does, the first merges remove far more of the manuscript than of any language: the first five merges each absorb 2 to 6 per cent of all symbols, where in Latin the first absorbs 2 per cent and the fifth 1. The EVA letters, in other words, are smaller than the script's natural units, which every transcriber has suspected and which chapter 5 discussed. It is the reason the solvers of the next chapter work on larger units than letters.

THE SHEET, NOT THE BINDING

Chapter 4 reported Layfield and Davis's argument, from the vocabulary, that the leaves were bound in the wrong order by someone who could not read them (Layfield and Davis, 2026). The claim can be tested with an independent measure. For every pair of pages with at least sixty words of running text we computed the similarity of their vocabularies, and compared pairs that stand in a physical relation—front and back of one leaf; the two leaves of one sheet; pages bound next to each other—with random pairs drawn from the same section and language, so that the comparison is not confounded by subject or by Currier's A and B.

Pages on one leaf are far more alike than random pages of their kind, and so are the two leaves of one sheet: the effect is nine to ten standard deviations from chance (figure 22.3). Pages that are merely bound next to each other, on different sheets, are not: the effect is one standard deviation, indistinguishable from chance, and pages two or three apart show nothing either. Pages in the same quire on different sheets show a faint signal, two standard deviations, which the Layfield–Davis account would attribute to the binder having kept some sheets roughly where they belonged.

This is, we think, the cleanest statement yet of what the text's order looks like from inside. The unit of composition is the sheet. The present sequence of sheets carries no vocabulary signal that a word-frequency measure can detect. The test does not say what the original order was; it does say that whatever determined the binding, it was not the text.

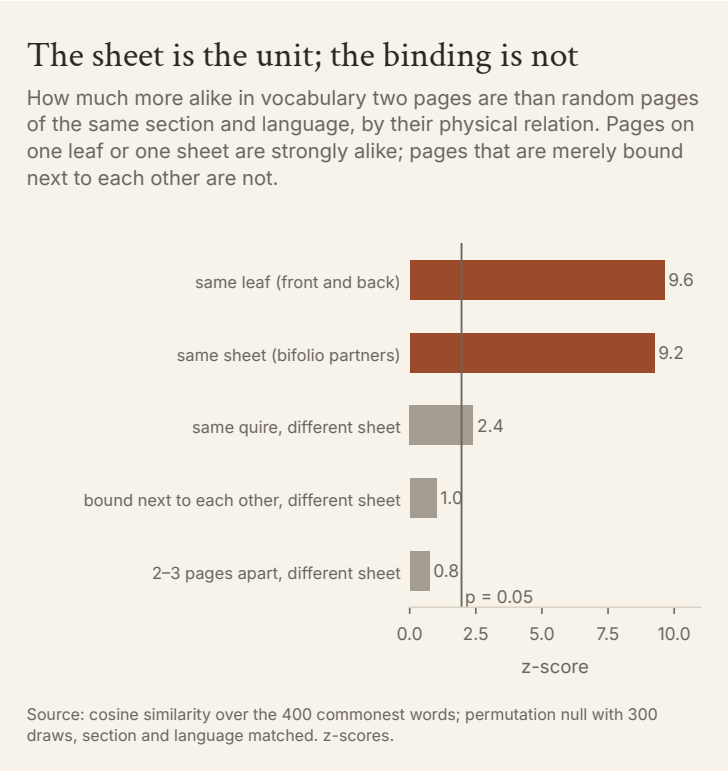


Figure 22.3 Vocabulary similarity by physical relation.

FORM AND FUNCTION

The second new result concerns the relation between how a word is spelled and how it behaves. In every language, words that differ by one letter are usually different words with different jobs: *cat* and *cut*, *et* and *ut*. Inflection produces exceptions, *domini* and *domino*, but they are a minority. In a text produced by a homophonic cipher the situation reverses: two words that differ by one symbol may be the same plaintext word written with a different homophone, and they will occur in the same contexts.

The test is simple. For the 300 commonest words of a text, measure how similar every pair is in spelling (one minus the edit distance, scaled) and how similar they are in use (the overlap of the words that precede and follow them), and correlate the two. Then, for each word, take its five nearest neighbours by use and ask how many are within one edit of it in spelling.

The results are in table 22.1. For twelve languages in plain form and nine in vowel-less or abjad form, the correlation between spelling and use lies between -0.03 and 0.09, and between 1 and 9 per cent of a word's nearest neighbours in use are its near-twins in spelling. For the manuscript the correlation is 0.15 and the share is 15 per cent; for the manuscript with its words shuffled, -0.03 and 3 per cent. Words that look alike behave alike, in the manuscript, to a degree no language in the sample approaches, including the vowel-less ones, in which near-twins are commonest.

Two of the surviving explanations of chapter 16 predict this. A homophonic or verbose cipher predicts it directly. The auto-copying generator predicts it too, since a copied word inherits the context of its source. A natural language in an unknown script does not predict it, and a plain nomenclator does not either. Together with the absence of repeated phrases, it is the strongest internal evidence the text offers that its common words come in sets of interchangeable variants.

How interchangeable? If every word with at least

Table 22.1 Form against function: spelling similarity versus similarity of use, 300 commonest words.

text	correlation	% near-twins	text	correlation	% near-twins
Voynich	**0.147**	**15.0**	Hebrew	0.055	5.7
Voynich, words shuffled	−0.025	3.3	Arabic	0.044	6.1
Latin	0.028	3.3	Syriac	0.013	7.4
Italian	0.032	6.9	Latin, no vowels	0.010	9.0
German	−0.027	3.4	Italian, no vowels	0.019	7.7
Czech	0.021	3.7	German, no vowels	−0.009	7.7
Hungarian	0.011	1.5	Czech, no vowels	0.026	9.2
Nahuatl	0.071	1.1	Hungarian, no vowels	0.016	6.9
Swahili	0.018	2.7	Greek	0.029	1.6
Quechua	0.085	1.6	Maori	0.042	3.1
Malagasy	0.090	2.7	Turkish	0.021	2.1
Finnish	0.017	1.3			

five occurrences is linked to every other within one edit, the 865 such words form ten connected families, and the largest holds four fifths of all occurrences. The same operation on Latin gives 600 families; on vowel-less Latin, 154; on Hebrew, 302; on vowel-less Swahili, 77. The manuscript's vocabulary is more densely interlinked than any attested text in the sample, and the texts that come closest are the ones written without vowels.

HIDDEN CLASSES

A last instrument from the decipherment literature is the hidden Markov model, which, given only a symbol sequence, learns a small number of hidden states and the probability of each symbol in each state. Trained on an alphabetic text with two states, it rediscovers vowels and consonants without being told what they are, because vowels and consonants alternate (Goldsmith and Xanthos, 2009; Knight and Yamada, 1999). We trained it on the glyph units, and, as a check, on Latin, Italian, Malagasy, Hebrew and Arabic.

On Latin, Italian and Malagasy the two states are the vowels and the consonants, exactly. On Hebrew and Arabic, written without vowels, the split is murkier, as it should be. On the manuscript the two states are **o**, **y**, **a**, **ch**, **sh** on one side and everything else on the other, and they alternate as strongly as vowels and consonants do in Latin (the model moves state four times out of five at every step). With three states the picture sharpens into something like a syllable. One state holds **o**, **e**, **y**, **a**: a nucleus. One holds **d**, **ch**, **sh**, **k**, **t**, **q**, **r**, **s**, **p** and the gallows-benches: onsets, which pass to the nucleus 95 times in 100. One holds **i**, **l**, **n**, **m**: codas, which follow the nucleus and precede the next onset. The model found, without supervision, the onset–nucleus–coda structure that the positional statistics of chapter 6 described by hand, and it put **i** and the **n**/**m** that end the **iin** groups in the coda, where a syllable-final element belongs.

This is one more way of saying the same thing: the script behaves as if each word were a short run of syllables of fixed shape. It does not say what the syllables are.

23

Brute Force

If the text is a European language in a cipher of the kind the fifteenth century used, then there is a key, the key is not large, and the manuscript is long enough to determine it. That last point is worth making exactly, because it changes what a failed search means. A homophonic key that maps fifty ciphertext units onto the twenty-three letters of a Latin alphabet and a space has about $50 \log_2 24 \approx 230$ bits of freedom. Latin written with that alphabet has a redundancy of roughly two and a half bits per character: the difference between the $\log_2 24$ bits a symbol could carry and the two bits or so it does. Dividing the one by the other gives Shannon's unicity distance, the length of ciphertext beyond which only one key produces text at all, and it comes to about a hundred characters (Shannon, 1949). The manuscript has a hundred and ninety thousand glyph units. For every cipher family in this chapter the same arithmetic holds: the key is small against the text by a factor of hundreds or thousands, and the obstacle, if there is a key, is finding it, not that the text is too short to pin it down. A solver that can find planted keys of the same size and fails on the manuscript is therefore evidence about the manuscript.

This chapter reports such searches. They are the brute-force part of the programme the user of this book might have expected to come first, and they come last because every one of them needs the rest: the units of chapter 5 and chapter 22, the held-out split and controls of chapter 20, and the 103-language reference of chapter 21.

THE SOLVER AND ITS CONTROLS

The search is for a many-to-one key: each ciphertext unit is assigned one plaintext letter, and several units may share a letter, which is what a homophonic or verbose cipher does. Four choices of unit were tried: the glyph units (29); the fillers of the slot grammar, so that **qo**, **aiin**, **edy** and the like are single units (50); 129 units found by merging the commonest adjacent pairs a hundred times, as a tokeniser does; and the 300 commonest whole words, with every other word as a single "rare" unit, for the hypothesis that each word is a letter. Each was tried with word spaces kept as a unit and with spaces removed. The plaintext languages were Latin, Italian, German, French and Czech, scored by a four-gram model trained on the whole Bible text of each, and vowel-less Latin, for the abbreviation and abjad hypotheses.

The solver is simulated annealing over keys, scored by the four-gram model with a penalty for keys that collapse the alphabet onto a few letters; it is a plain version of the method that solved the Zodiac killer's 340-character cipher in 2020 (Oranchak et al., 2024). Before it was shown the manuscript it was given Latin enciphered with keys we chose, eighty units with up to four homophones per letter, fitted on 24,000 words and verified on the remaining 10,000. It recovered the planted key completely with spaces kept and 98 per cent of it with spaces removed, and the recovered text scored on held-out words within 0.03 nats per character of the language's own text with spaces and 0.08 without. That margin is the standard a search on the manuscript has to meet. The control also marks the method's limit: vowel-less Latin with its spaces removed as well defeats the solver (49 per cent of the key recovered, and a decoded text the language model likes better than the original), so that configuration was not searched; what the solver cannot find in a planted text it cannot be credited with failing to find in the manuscript.

WHAT THE SEARCHES FOUND

Table 23.1 gives every search. Two columns matter. The first is the gap between the score of the decoded held-out text and the score of the language's own text, in nats per character: for the planted keys it was 0.02 to 0.08; for a random key it is around nine. The second is the share of decoded held-out words that are words of the language, against the share a random key produces.

For Latin, Italian, German, French and Czech under glyph or slot units with spaces kept, the gap is between 1.8 and 2.6 nats, the keys found use nine to seventeen letters with half to three quarters of all units assigned to the three commonest, and the "decoded" text is of the kind the samples show: *eneeg rgen ege gerin*. The dictionary rates, 10 to 33 per cent against 1 to 3 for uniformly random keys, are made of short words—*et*, *in*, *es*, *ne*—and of the common letters the key collapsed onto. Removing spaces narrows the gap, to 0.7 for Latin and 1.1 for German, for the uninteresting reason that a language model with no spaces to predict is easier to satisfy with a monotonous string. The merged-pair units did worse than the glyphs, and the word-as-letter search did worst of all: with 300 words as units the solver could do no better than assign nearly all of them to

Table 23.1 Homophonic searches on the manuscript. Gap: the language's own four-gram score minus that of the decoded held-out text, in nats per character (planted-key controls 0.02–0.08; random keys about 9). Words: share of decoded held-out words in the language's dictionary, found key / random key, and the same for words of four letters or more. Permuted: the found key's dictionary share when its letters are reassigned among the units at random (mean of twenty). Letters used: distinct plaintext letters in the found key, and in brackets the share of units assigned to the three commonest.

language	units	spaces	gap	words %	random %	≥4 letters %	permuted %	letters used
Latin	glyphs	yes	1.88	17	2	10	5	13 (0.52)
Latin	glyphs	no	0.70	18	2	11	6	10 (0.66)
Latin	slot fillers	yes	1.77	28	2	15	12	12 (0.48)
Latin	slot fillers	no	0.69	26	2	2	14	10 (0.66)
Latin	100 merges	yes	2.49	64	8	6	41	11 (0.72)
Latin	300 words	no	2.54	24	9	0	35	14 (0.62)
Italian	glyphs	yes	1.96	26	2	10	10	12 (0.55)
Italian	slot fillers	yes	2.04	33	2	16	13	11 (0.62)
Italian	slot fillers	no	1.07	24	2	6	14	12 (0.54)
Italian	300 words	no	3.10	100	41	0	99	17 (0.53)
German	glyphs	yes	2.58	10	1	5	3	12 (0.48)
German	slot fillers	yes	2.16	27	1	19	6	9 (0.68)
German	slot fillers	no	1.09	15	1	8	6	9 (0.76)
German	300 words	no	3.48	53	10	0	23	13 (0.67)
French	glyphs	yes	2.42	18	1	10	6	9 (0.62)
French	slot fillers	yes	2.38	23	2	15	10	10 (0.54)
Czech	slot fillers	yes	2.30	24	3	8	7	17 (0.48)
Latin, no vowels	glyphs	yes	0.41	66	14	53	39	8 (0.62)
Latin, no vowels	slot fillers	yes	0.34	82	31	63	57	10 (0.74)
control: Latin, planted key	80 units	yes	0.03	96	2	–	–	key 100 % recovered
control: Latin, planted key	80 units	no	0.08	0	0	–	–	key 98 % recovered
control: Latin, no vowels, planted key	80 units	yes	0.02	98	55	–	–	key 98 % recovered
control: Latin, no vowels, planted key	80 units	no	-0.04	0	0	–	–	key 49 % recovered

a and *n*, a gap of two and a half to three and a half nats, which is the result one gets when there is no structure of the assumed kind to find.

A uniformly random key is too weak a null for the dictionary column, because the solver's keys are not uniform: they pile units onto frequent letters, and frequent letters make dictionary words by accident. So for every search we also took the key the solver found, kept its multiset of letters, and reassigned them among the units at random, twenty times. Those permuted keys reach 3 to 14 per cent dictionary words for the five languages under glyph or slot units, against 10 to 33 for the keys the solver found, and 41 per cent for the merged-pair units (where the found key reaches 64): most of what the found keys achieve, a permutation of them achieves too.

Vowel-less Latin is the one case that needs a second look, and the one for which the control was run separately. Latin with its vowels removed and enciphered with a planted key was recovered 98 per cent, with a gap of 0.02 and 98 per cent dictionary words against 55 for random keys: a vowel-less dictionary of consonant skeletons is permissive, and the random rate is high. On the manuscript the best vowel-less key gives a gap of 0.34, seventeen times the control's, and 82 per cent dictionary words, against 57 for its own permutations and 66 for the best of them; counting only words of four letters or more, 63 against 23. The found key does better than its per-

mutations, by more than any other search in the table. It does not meet the rule of chapter 20: its score does not transfer to the held-out sheets within the control margin, and the text it produces, *tstts rsts dst strns*, is not Latin by any scorer that did not take part in the search. What it shows is that the manuscript's glyph sequence, mapped onto a handful of consonants, has the bigram texture of a consonant skeleton, which is the abjad proximity of chapter 21 seen from the other side. We record it as the most interesting negative in the table and as the configuration a better solver should start from. (A note added after Part VI: when the searches were repeated with the stronger penalty against collapsing the alphabet that the solver of chapter 26 needed to pass its controls, the plain-language gaps were unchanged, 2.0 to 2.5 nats, and the vowel-less Latin gap widened from 0.34 to 0.72 as the key spread over the alphabet. The near miss was the collapse.)

TRANSPOSITION AND SHIFTING ALPHABETS

The homophonic search assumes that the units are in plaintext order. Two other families would defeat it: a transposition, in which the units are written out in a scrambled but regular order, and a polyalphabetic or position-keyed substitution, in which the key changes with position in the line. The Zodiac-340 solution began from a diagnostic for the first, a count of repeated bigrams when the text is read at every period, which peaked at the cipher's period of nineteen (Oranchak et al., 2024). On the manuscript that count is flat: the best period exceeds the median over all periods by four hundredths of one per cent, in the text and in every surrogate alike. The periodic coincidence rate peaks at six units, which is the mean word length plus its space, and does so in every surrogate that keeps the words. There is no period.

For the second family we counted, across all pairs of lines, windows of six to nine units with the same pattern of repetition but no unit in common, the isomorphs of Enigma cryptanalysis, which a position-keyed alphabet would place at matched line positions. Two and a half per cent of such pairs fall at matched positions in the text; 2.4 per cent in the text with words shuffled within lines. Consecutive lines coincide at matched positions 10.5 per cent of the time against 9.4 for random line pairs, and the shuffle keeps most of that excess, which means neighbouring lines share words, as chapter 22 found, not positions. Nothing in the text is keyed to its position in the line.

NULLS AND ANAGRAMS

Two older proposals were tested the same way. If one or two glyph units are nulls, meaningless insertions, then deleting them should lift the conditional entropy of the rest into the range of attested languages. Deleting any single unit, or any of the five commonest pairs, lifts it from 2.26 to at most 2.40 bits, against an attested minimum of 2.51; the same deletion from twelve languages moves their entropy by a tenth of a bit. If each word is a set of letters written in a fixed canonical order, an anagram cipher with sorting, then nearly every word should be monotone under one global ranking of the units.

Forty-four per cent are, against 9 per cent for the same words with their letters shuffled; for twelve languages the share runs from 10 to 32 per cent. The manuscript's spelling is more predictable from its letter set than any language's, which the slot grammar already said, and it is far from fully predictable, which rules out the sorted form of the hypothesis.

HOW MUCH IS THERE TO FIND?

The searches above all looked for a reading in the order of the units within words. Two further measurements ask how much information the order of the words carries, and they bear on every cipher hypothesis at once.

The first is a law. Across more than a thousand languages, the entropy of single words and the entropy rate of the word sequence lie on a line: a text with a large vocabulary is correspondingly harder to predict word by word, and knowing the preceding words removes about three bits of the cost (Bentz et al., 2017). We refitted the relation on 102 of our equal-length samples with a compression estimate of the rate (the Japanese sample is written without word spaces and was left out), where it holds with a correlation of 0.98, and placed the manuscript on it (figure 23.1). It lies 3.7 standard deviations above the line: harder to predict than a language with its vocabulary would be. More telling is the gain from word order alone, the number of bits per word saved by knowing the sequence rather than the shuffled bag: 0.42 for the manuscript, against 1.34 for the lowest of the 102 languages and a median of 2.39. No language comes within a factor of three of it.

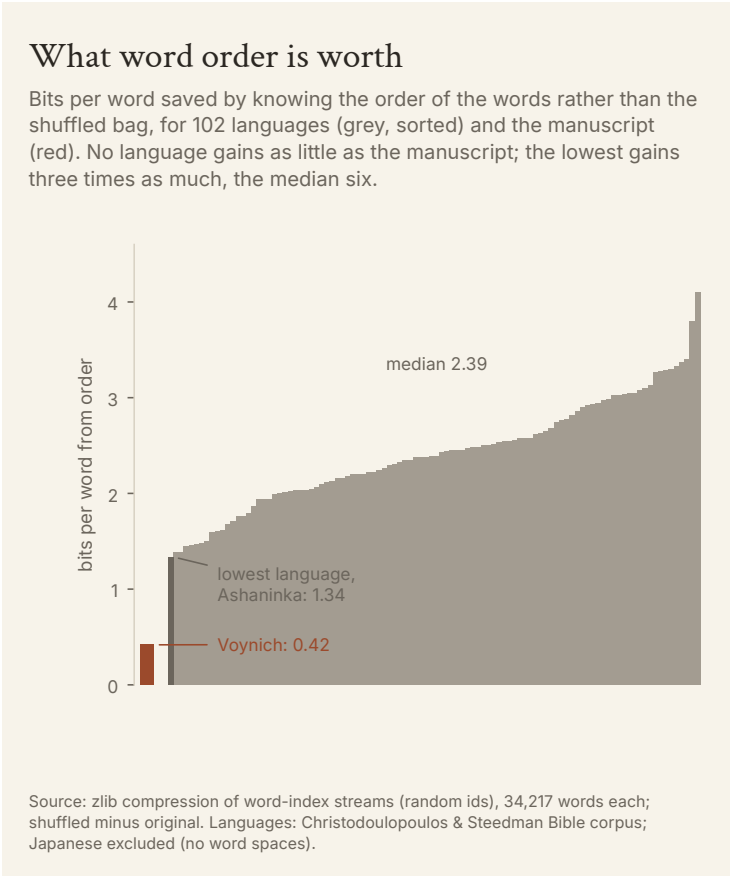


Figure 23.1 The gain from word order, 103 languages and the manuscript.

The second is a held-out test of the same thing. A model that predicts each word from the one before, fitted on the training quires, predicts the held-out quires *worse* than a model that ignores the preceding word, by 0.15 bits per word; in eight languages the preceding word helps by 0.3 to 1.2 bits. At the level of the glyph the picture is the same in miniature (figure 23.2): nearly all the predictability is reached by a context of two units, and lengthening it to a whole word gains a fifth of a bit, where in Latin it gains 1.4.

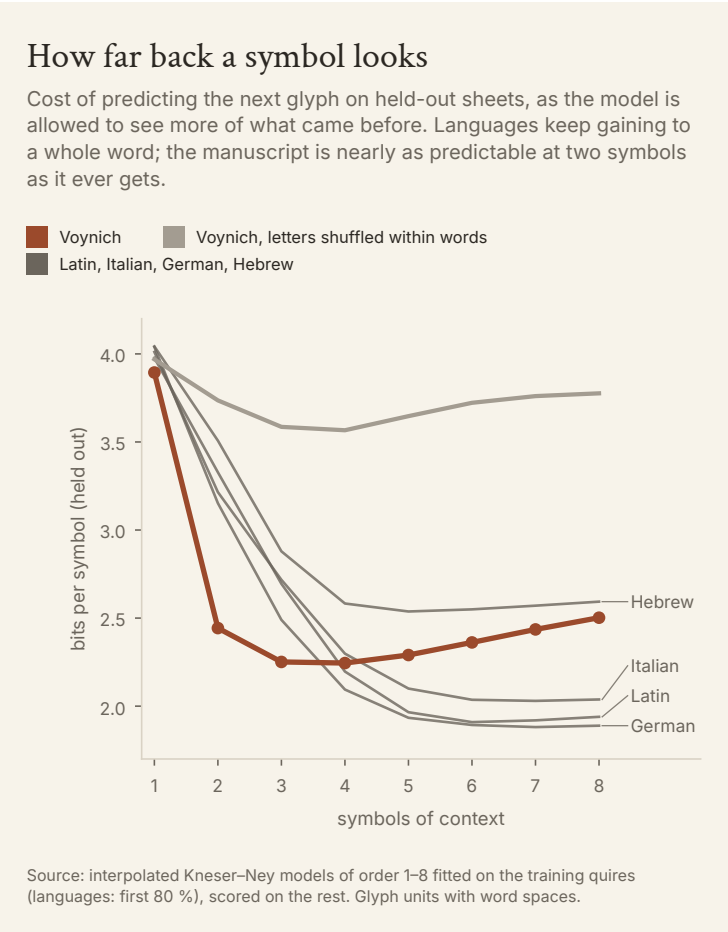


Figure 23.2 Held-out predictability against context length.

These two results constrain the cipher hypotheses more than any search does. In a language, the order of the words carries information, and a substitution cipher of any kind preserves it. The manuscript's word order carries less than any language's. So either the cipher destroys word order, which a substitution does not and a transposition, excluded above, would; or the plaintext is not running prose but something like a list, whose items can be reordered at little cost; or there is no plaintext. A solver looking for prose under a substitution is, on this evidence, looking for something the text does not contain, however good its language model, and that is the main thing this chapter has to say about brute force.

24

Pictures as Data

The pictures have had a harder time than the text. The text can be counted; the pictures have mostly been looked at, and what one person sees in a drawing of a plant or a bathing woman is not easily checked by another. Part III described the pictures. This chapter treats them as the text has been treated: as a source of measurements, with the measurement stated, the null comparison stated, and the limits of a 2,000-pixel photograph admitted.

WHERE THE PAINT IS

The simplest measurement is of paint. Every page photograph was classified pixel by pixel into green, red-brown, yellow, blue, ink and parchment by thresholds in hue, saturation and value, chosen on a dozen pages by eye and then applied to all 193 ordinary pages (the fold-outs were excluded because their photographs include the backing). The result, figure 24.1, is a map of the painter's work in bound order.

Green is the herbal pigment: a median of 4 per cent of the surface of a herbal page, against a fraction of a per cent elsewhere, with the bathing pages the only other section in which it is regular. Blue, which has been the subject of more speculation than any other pigment because of its cost, covers more than 0.3 per cent of the surface on 22 pages, 16 of them herbal and the rest scattered through the astronomical, bathing and pharmaceutical sections; it never dominates a page. Three herbal pages carry effectively no paint at all (f21r, f30v, f56r), and the recipes section, with its stars, carries only the ink of the stars themselves.

The hypothesis the map was built to test is that the painting was done in campaigns organised by gathering, so that paint coverage should vary by quire more than by scribe or by Currier language within the herbal section. It does not, at this resolution. Within the 120 herbal pages, the quire accounts for 8 per cent of the variance in green coverage, the scribe for 4 per cent and the language for 2 per cent, and none of the three is distinguishable from a random assignment of the same coverages to pages (permutation p = 0.40, 0.23 and 0.28). Pages differ in how much is painted because plants differ in how much leaf they have, and that swamps whatever the painter's schedule contributed. The registry records the hypothesis as not supported at 2,000 pixels; a multispectral or higher-resolution campaign could still find what this one cannot, and the data layer is there for it.

THE ZODIAC AS A TABLE

Figure 24.2 lays the twelve zodiac wheels out as a table rather than a picture, because the questions that can be answered about them are tabular. For each wheel: the month written beside the emblem by a later hand, the count of figures, the count of label words, and the share of labels beginning with the two commonest openings. The figure counts are the familiar ones: thirty per sign, except that Aries and Taurus are each divided across two wheels of fifteen and Pisces has twenty-nine. The month names run from March to December with no January or February, which is the first and plainest fact about the sequence: two signs, Capricorn and Aquarius, are missing, and so is the leaf that would have carried them.

The labels are the more interesting column. Of the 345 label words beside the figures, 281 are distinct and 37 of those recur on more than one wheel, which rules out the simplest reading in which each label is the unique name of one star or one degree. And the labels drift. On Pisces and the two Aries wheels, 44 to 71 per cent of labels begin with **ot-**; from Cancer onward the share falls to 5 to 18 per cent and labels beginning **ok-** take over, reaching half on Gemini. Whatever the first syllable encodes, it is not a property of the individual figure, since the figures are drawn alike from wheel to wheel; it tracks the wheel, or the order in which the wheels were written. This is the kind of result a reader should expect from labels if the opening syllable is a marker rather than a letter, as the slot grammar of chapter 6 suggests.

THE ROSETTES, NUMBERED

The rosettes sheet (figure 24.3) is the picture about which most has been claimed and least can be measured. The plate numbers the nine circles for reference and describes only what is drawn in each: stars of blue rays in two, a field of dots in one, a ring of towers or pavilions at the centre, walls with forked battlements in the ninth, and in the corner the small circle divided into a half and two quarters that is the T-O schema of the medieval world map, drawn in the same way on a thousand other pages of the period. The orientation is Yale's; which way the sheet was meant to be read is unknown, and the numbering carries no claim about it.

What can be measured is the writing, and it gave

the one result of this chapter that we did not expect. The sheet carries 536 words: 152 in short labels, 37 in a paragraph, and 337 in rings running round the circles. Chapter 7 showed that labels differ from paragraph text in a fixed way: fewer than 3 per cent begin with **q**, against 15 per cent of paragraph words, and more than half begin with **o**, against 20 per cent. The rosette rings are written like labels, not like paragraphs: 1.8 per cent of their words begin with **q** and 61 per cent with **o**. The sheet's one paragraph, by contrast, has 11 per cent **q** and 46 per cent **o**, as paragraph text on these pages does.

THE REGISTER OF RINGS

That prompted a wider test (table 24.1). Across the whole book, text written in rings round diagrams amounts to 2,349 words on the astronomical, zodiac and cosmological pages, and radial text to 346 more. Pooled, the ring text begins with **q** 1.9 per cent of the time and with **o** 40 per cent; radial text, 5 and 35 per cent; labels, 1 and 54 per cent; paragraphs, 15.5 and 20 per cent. On the seven pages that carry at least thirty words of both ring and paragraph text, the paragraph text has on average 6.5 points more **q**-initial words than the ring text on the same page, a difference that never arose in five thousand reshufflings of the words between the two.

The registry had predicted the opposite: that ring text, which runs on continuously for dozens of words, would be paragraph text bent into a circle. It is not. Whatever **q** marks, and whatever the **o**-initial preference of labels reflects, both go with the paragraph as a layout rather than with the length of the text, and a scribe writing round a circle wrote as he did beside a figure. Of the explanations in chapter 7, this favours the orthographic one: **q** is a feature of lines in paragraphs, and words outside paragraphs, long or short, do without it. It is also a warning to any solver that treats the ring text as a sample of the same cipher as the paragraphs; on this evidence it is the same system in a different mode, and the two should be fitted separately.

Beyond that the sheet has resisted every quantitative question we have put to it, and the honest report is that the castle and the T-O map are the only two elements on it that can be matched to anything outside the manuscript without interpretation.

WHAT THE PICTURES HAVE BEEN COMPARED WITH

Comparison with other manuscripts is where the pictures have been most studied and least measured, and a short account of the state of that work belongs here, because it sets the standard the next comparisons should meet. The established points are few. The zodiac emblems are of a Central European type: the crossbow-carrying Sagittarius is found only in German-speaking regions, and Zandbergen's comparison of more than a hundred zodiac cycles found its nearest relatives in South German manuscripts (Zandbergen, 1996–2026). The bathing section has been compared since the 1990s with the illustrated copies of Peter of Eboli's *De balneis Puteolanis*, in which, as in the manuscript, the water is painted green. The herbal drawings belong, loosely, to the family of the north Italian "alchemical herbals" of the late fourteenth and fifteenth centuries, and one

drawing, on f35v, closely resembles a plant in a Paris copy of the *Tractatus de herbis* (BnF lat. 6823, f. 60r), though Zandbergen judges it inspired by a similar image rather than by that one. The T-O circle on the rosettes sheet is a standard cartographic schema. None of this identifies a model; all of it points to the same corner of Europe and the same half-century as the parchment.

Beyond those points the comparisons are proposals. The most rigorous is Marco Ponzi's, who scored 131 zodiac cycles on twenty features and found the manuscript's cycle closest to manuscripts from Poland, Bavaria and Styria, with a Gemini pose that matches only one of them (Ponzi, 2016); Koen Gheuens traced that pose, a couple with crossed hands, to a marriage image used in the Hagenau workshop of Diebold Lauber in the 1430s and 1440s (Pelling, 2018). The rosettes sheet has been read as a map of the elements, as a plan of a particular city, as a cosmological diagram and as a clock (Pelling, 2006; Wastl and Feger, 2014); none of these readings has been tested against a competing one. The ten folios imaged multispectrally in 2014 and published in 2024 made the f1r inscriptions legible and showed that several lines on f116v once read as ink are stains, but revealed no hidden text (Davis, 2024).

What distinguishes Ponzi's comparison from the others is that it is a table, and could be re-scored, extended, and tested: one can ask how often a randomly chosen German cycle sits as close to its nearest neighbour as the manuscript's does. We did not do this, because his per-cycle scores are not published as data, but it is the first thing we would do with the pictures next, and chapter 30 lists it with the other candidates.

THE SCRIPT AS A MACHINE

The last figure of this chapter is of the script itself. Figure 24.4 draws the glyph units as circles, sized by frequency and placed left to right by their mean position in the word, and connects each to the units that follow it at least one time in six. The result is not a network in the social sense but a flow diagram, and it reads as one: the openers **q**, **o**, **sh**, **ch** and the gallows on the left; **e**, **a**, **k**, **d** and **s** in the middle; **y**, **l**, **r**, **n**, **m** on the right, with nothing after them. Almost every arrow runs rightward. The few that run back, **a** to **ch** for instance, mark the start of a second syllable within a word.

The diagram is a picture of the slot grammar of chapter 6 and of the three-state model of chapter 22, drawn from the transitions alone. It is included because it answers, more directly than a table of positions can, the question a reader most often asks on first seeing the script: why does it look so regular? It looks regular because it is a small machine with a fixed direction, and the drawing is the machine.

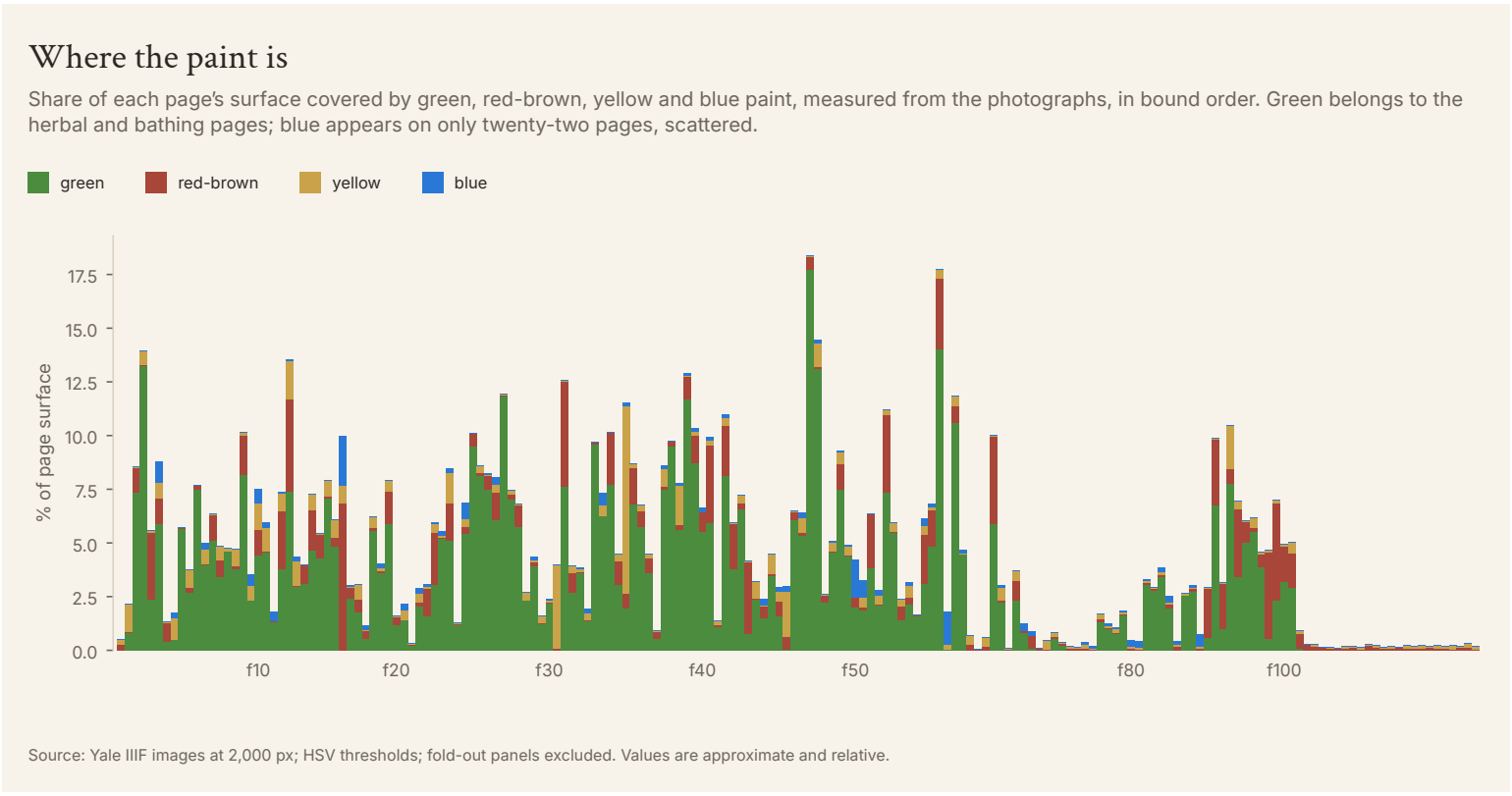


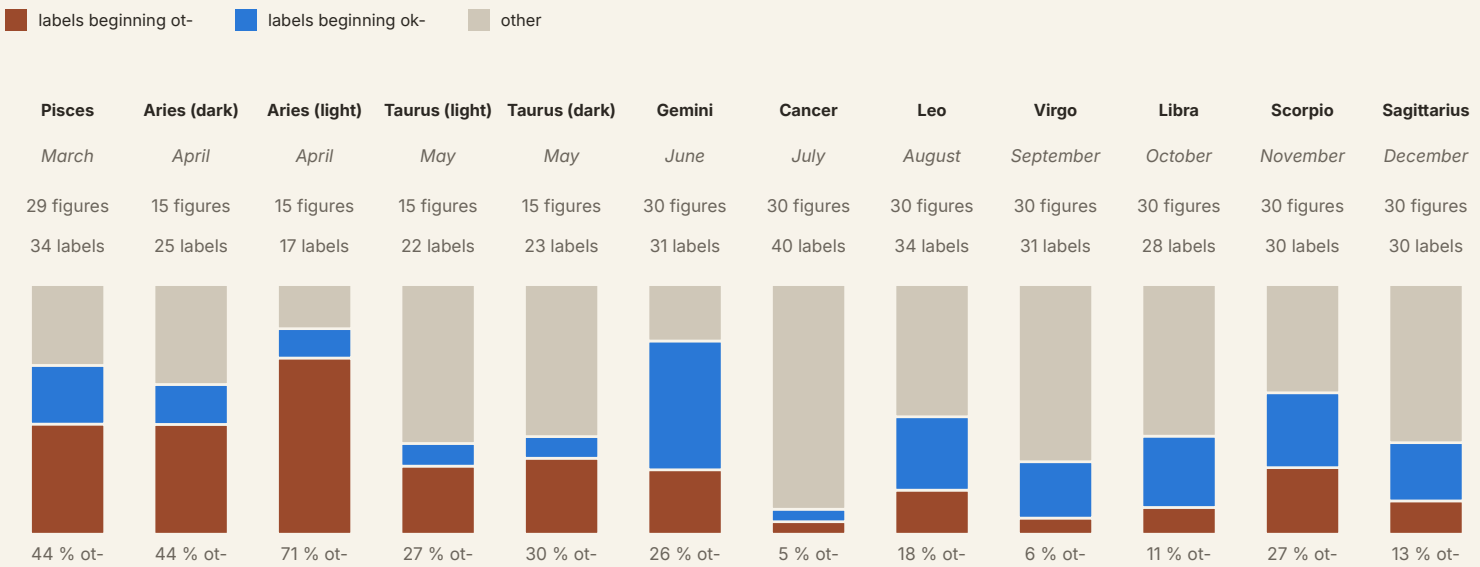
Figure 24.1 Paint coverage by page, in bound order.

Table 24.1 The register of each kind of text: share of words beginning with **q**, with **o**, with **ch** or **sh**, and ending in **dy**; distinct words per hundred.

section	text	words	**q**- %	**o**- %	**ch**/**sh**- %	-**dy** %	distinct %
all	paragraphs	34,217	15.5	20.4	24.9	18.0	22
all	rings	2,349	1.9	39.6	21.9	12.4	47
all	radial	346	4.9	35.3	14.7	15.0	80
all	labels	1,165	1.0	54.1	8.2	14.8	72
astronomical	paragraphs	276	7.2	21.7	21.4	10.9	80
astronomical	rings	291	2.1	33.3	27.1	13.1	76
zodiac	rings	951	1.5	40.6	26.4	11.3	58
zodiac	labels	345	0.3	67.2	8.4	15.1	81
cosmological (incl. rosettes)	paragraphs	692	12.3	21.7	22.1	16.2	61
cosmological (incl. rosettes)	rings	1,107	2.3	40.5	16.6	13.3	52
rosettes	paragraph	37	10.8	45.9	18.9	16.2	95
rosettes	rings	337	1.8	61.4	8.6	14.5	64
rosettes	labels	152	0.7	63.2	3.9	28.9	91

The zodiac, sign by sign

For each of the twelve wheels: the month written beside the emblem in a later hand, the number of figures, the number of label words beside them, and how those labels begin. Labels on the early signs mostly open with ot-; on the later signs with ok-.



Source: ZL3b zodiac labels (type Lz); figure counts after Zandbergen. Aries and Taurus are each split over two wheels of fifteen.

Figure 24.2 The twelve zodiac wheels as a table.

The rosettes sheet, numbered

Folios 85v–86r unfolded, in the orientation of Yale's photograph. The nine circles are numbered for reference in the text; the list describes only what is visibly drawn in each. Which way up the sheet was meant to be read is not known.



1. ring of scale- or tube-like cells round an oval core with a blue band

2. star of blue rays inside a scalloped rim

3. a field of dots or stars inside a wavy rim

4. pale rays and a ring of text; paths lead to 1 and 7

5. the central rosette: a ring of small towers or pavilions and a ring of figures round a patterned core

6. dotted scales with a ring of text

7. scaled pattern with a small inner ring

8. star of blue rays, like 2

9. walls and towers with forked battlements at the rim

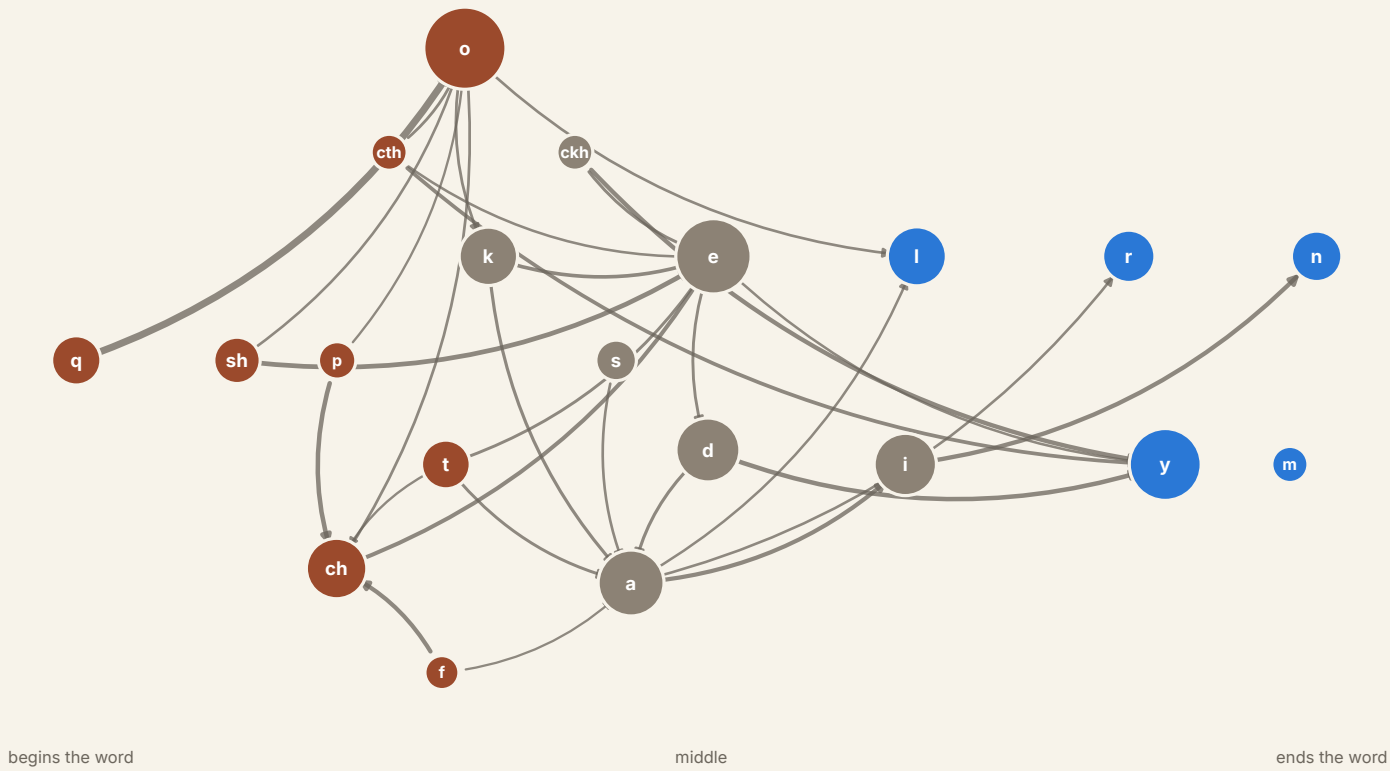
T. a small circle divided into a half and two quarters: the T-O schema of the medieval world map

Image: Beinecke Rare Book & Manuscript Library, Yale University, Beinecke MS 408. Numbering: the authors.

Figure 24.3 The rosettes sheet, numbered.

The grammar of the script, as a map

Each circle is a symbol, sized by frequency and placed left to right by where it usually sits in a word. An arrow from A to B means that when A is written, B follows it at least one time in six. The script reads as a left-to-right machine: openers, a gallows, a middle, an ending.



Source: ZL3b running text, glyph units; transitions with P(B|A) ≥ 0.16 shown. Vertical placement spreads symbols for legibility and carries no meaning.

Figure 24.4 Glyph transitions as a flow diagram.

Part VI

The Procedure

25

Five Hands, One Procedure

Five people wrote this book, and their text cannot be told apart. Chapter 8 found that once subject and language are held constant, the scribe accounts for less than one per cent of the vocabulary. Whatever produced the text was therefore not one person's improvisation. It was a procedure that five people carried out the same way, or a text that five people copied from one source. Either way it was mechanical, and that is the most useful thing this book knows about it, because a mechanical procedure leaves fingerprints that a free composition does not.

This chapter does three things with that. It asks whether the hands show any trace of executing the procedure individually, which would separate execution from copying. It sets out what procedures of the period could have done the job, drawing on a survey of the cipher practice of the chanceries between 1380 and 1520. And it builds six of those procedures as simulators, fits each to the manuscript, and asks which of its measured properties each one reproduces.

DO THE HANDS DIFFER?

If each scribe was enciphering for himself, choosing among the variants a table allowed, his choices should carry a bias of his own: a preference for **ch** over **sh** here, for **-aiin** over **-ain** there. If the scribes were copying, the choices were made once, upstream, and the hands should agree on them exactly while differing in the errors copyists make. The one place in the manuscript where this can be tested cleanly is the herbal section in Currier's language B, where three scribes, Davis's 2, 3 and 5, wrote 3,316

words on 32 pages of the same kind.

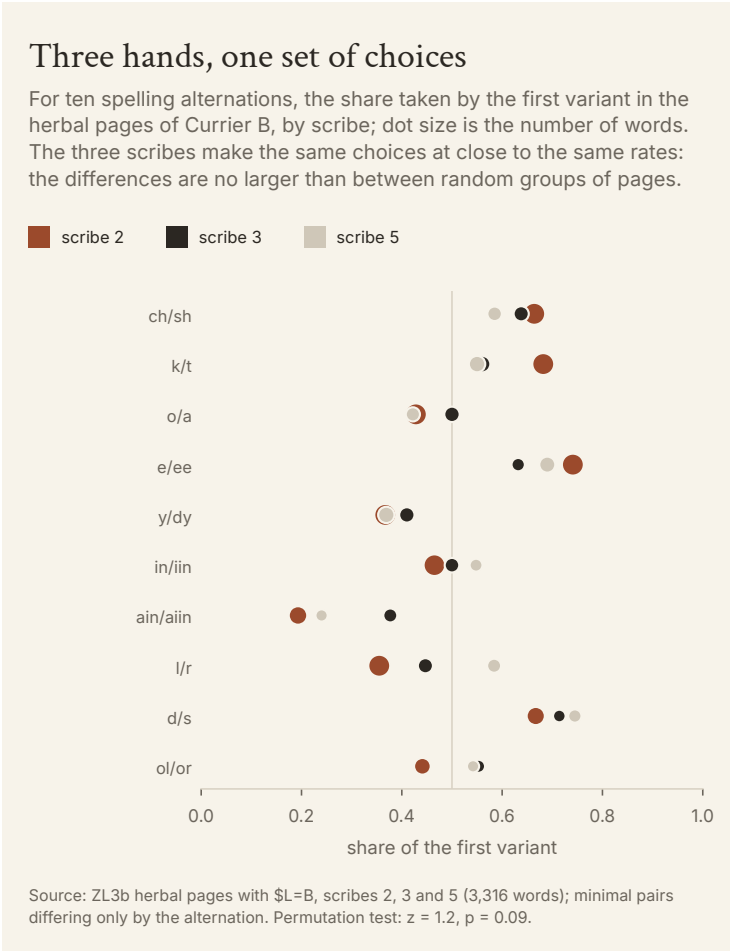


Figure 25.1 Variant choice by scribe in the herbal pages of language B.

For ten alternations between variant spellings we took every pair of words that differ only by that alternation and counted, per scribe, how often each member was chosen (figure 25.1). The null is the same count with the scribe labels shuffled across pages within the section and language, so that a difference between pages of different subject cannot pass for a difference between hands. The three hands agree: the summed divergence of their choices is 1.2 standard deviations above what random groupings of the same pages give (p = 0.09). Scribe 2 chooses **ch** over **sh** 66 per cent of the time, scribe 3 64, scribe 5 59; **y** over **dy** 37, 41 and 37; **d** over **s** 67, 71 and 75. The share of near-twin substitutions that are visual confusions (**ch**/**sh**, **k**/**t**, **a**/**o** and the like) is 17, 17 and 18 per cent. The same statistics applied to one scribe writing two subjects, scribe 1 in the herbal and in the pharmacy, give 19, 17 and 3 standard deviations: the subject moves the text by an order of magnitude more than the hand does.

Two of the three hands are represented by only five hundred words, so a small bias could hide. But a bias of the size people show when choosing among homophones by hand, which in the Zodiac ciphers appears as cycling and avoidance of the symbol just used, is not there. The glyph-transition profiles of the three hands do differ slightly (6.9 standard deviations), which is what a copyist's or an executor's habits of ligature and spacing would produce. The evidence is consistent with a fixed table executed by rote or with copying from one exemplar; it is not consistent with five people composing freely.

WHAT PROCEDURES WERE AVAILABLE

The chanceries of Italy had been encoding text by table since the fourteenth century, and the surviving keys show what a table could look like in the decades the parchment was made (Pelling, 2016). Gabriele de Lavinde's ledger of 1379 gives each letter one sign, adds a dozen signs for whole words, and a few nulls. The Mantua alphabet of 1401 gives each vowel four signs, so that the vowels, which a cryptanalyst finds first, are hidden. The Tranchedino ledger of Milan, around 1450, gives vowels three signs and consonants two, has signs for doubled letters and for the common abbreviations, more nulls, a nomenclator of over eighty entries for names and common words, and in the Mantuan key of the same decade a table of syllables, *ab*, *ac*, *ad*, each with its own sign. The rules Cicco Simonetta wrote at Pavia in 1474 for breaking such ciphers assume that word divisions are visible and attack the word endings, which is to say that the designers of the next generation of keys knew exactly what to hide. Trithemius, writing just after our window closes, replaced each letter by a whole Latin word from a column of a table, so that the ciphertext read as a prayer; the third book of his *Polygraphia* does the same with invented words that look like inflections of a single stem (Hermes, 2022).

Every one of these is a procedure that a clerk could carry out from a sheet, and none of them, applied to Latin, gives Voynichese. A one-sign-per-letter key relabels the plaintext and keeps its entropy; chapter 21 excluded it against 103 languages. Vowel homophones raise the entropy rather than lowering it. A nomenclator makes a handful of very frequent short tokens, which the manuscript lacks. Hermes measured the *Polygraphia*'s word-for-letter tables against the manuscript and found the conditional entropy in the right range, but seven times too many minimal pairs per unit and three times too many immediate repetitions; and in a word-for-letter table, words that look alike encode *different* letters, so the coupling of spelling to use that chapter 22 found would come out negative, not positive (Hermes, 2022). The nearest attested relative is the Tranchedino family, because it mixes units of three sizes, letter, syllable and word, under one key; and the one proposal built expressly to produce Voynich-like text, Greshko's Naibbe cipher, is a Tranchedino-like table of letter and letter-pair units with six forms each, chosen by cards (Greshko, 2025). Its author concedes that it produces no dependency across word spaces, a singleton share of 41 per cent against the manuscript's 70, and no line effects.

SIX PROCEDURES, SIMULATED

The way to compare procedures with a text is to let each procedure write the text. We built six simulators, each of which writes into the manuscript's own layout, page by page and line by line with the real counts of words per line, so that every statistic of this book can be computed on the result exactly as on the original. Each simulator has a few free parameters, and those were fitted on four statistics of the training quires, the conditional entropy, mean word length, the share of distinct words and the share of words used once, so that the comparison is made on statistics the fitting did not see.

The six are: words drawn independently from the slot grammar of chapter 6; Timm and Schinner's self-copying, in which each word is a copy of a recent word with a small change (Timm, 2014; Timm and Schinner, 2020); Rugg's table and grille, a table of syllables read through a sliding card (Rugg, 2004; Rugg and Taylor, 2017); an affix-table cipher of the Naibbe kind, in which each Latin letter or letter pair is written as a prefix and a suffix drawn from tables, plaintext spaces dropped; a Tranchedino-type homophonic cipher of Latin with nulls, its stream cut into words of the manuscript's lengths; and a Trithemius-type word-for-letter table. For the three cipher procedures the plaintext is the Latin Bible, and the affix-table cipher was also run on vowel-less Latin. Twenty corpora were written by each.

Figure 25.2 is the result, laid out as a grid of procedures against ten held-out statistics, with three real languages written into the same pages for comparison. No procedure reproduces more than two of the ten. The pattern of what each reproduces is more informative than the count.

The affix-table cipher, with sixty prefixes and sixty suffixes per unit, comes closest on the statistics it was fitted to, and it reproduces the one property that no language has: it makes word order worthless. Its word-order gain is zero against the manuscript's 0.40, and the previous word makes the next harder to predict on held-out pages, as in the manuscript. This is worth stating plainly, because chapter 23 presented the weak word order as evidence against a substitution cipher of prose. It is evidence against a cipher in which the written word is the plaintext word. A cipher in which each plaintext letter becomes a word, and each letter has dozens of forms, hides the order of the plaintext behind the randomness of the table, and the manuscript's word order is, if anything, slightly *more* informative than such a cipher's. But the same procedure fails three things a language has and the manuscript has: it has no topical locality (adjacent pages are no more alike than random pages, z = 1.9, where the manuscript gives 16 and languages 10 to 20), no coupling of spelling to use (0.04 against 0.15), and no dependency across word edges (0.01 bits against 0.21).

Self-copying reproduces the near-twin rate and the share of near-twins among a word's nearest neighbours in use, which is what copying with small changes does; it over-produces order (2.9 bits per word from word order, because each word is predictable from the last) and, in the form simulated, has no page locality. The homophonic stream cut into words is the only procedure that reproduces the dependency across word edges, 0.22 bits, because its word boundaries fall inside a continuous dependent stream; it fails nearly everything else. The word-for-letter table, the grille and the independent slot grammar fail across the board. No procedure reproduces the line-initial effect, which none of them has by construction, and which the manuscript shares with no language either: it is a property of the page, not of the text, and any procedure that produced the manuscript did it at the desk, line by line.

WHAT THE FINGERPRINTS SAY

Put the three columns of evidence side by side. The hands are interchangeable, as a table executed by

rote or a copied exemplar would make them. The word-level statistics, weak order, high singleton share, low entropy, are those of a verbose cipher with large tables. The page-level statistics, topical locality and the coupling of spelling to use, are those of a text whose words mean something and recur where their subject recurs. No single procedure in our set has all three, and the two halves pull against each other: the larger the table, the better the word statistics and the worse the locality, because a word that can be written sixty ways cannot recur.

The manuscript's words do recur, by subject, across adjacent pages and across the two leaves of a sheet. So whatever the procedure was, its output words were not as free as sixty forms per letter would make them. Either the tables were small and something else lowered the entropy, or the units were larger than letters, syllables or whole words with a few forms each, so that the same plaintext item tends to come out the same way and the vocabulary keeps its topical shape. That is a Tranchedino-type mixed table rather than a Naibbe-type letter table, and it is the configuration the solver of the next chapter was built to search.

Fingerprints: what each procedure reproduces

Ten held-out statistics of the manuscript (left column) against six mechanical procedures fitted on the training quires and three languages laid out in the same pages. A filled cell means the value is within a quarter of the manuscript's (or, for z-scores, within 2); grey means it is not. No procedure reproduces more than two.

matches the manuscript does not

	procedures						languages		
	affix-table cipher	**homophones, respaced**	**word for letter**	**self-copying**	**table and grille**	**slot grammar**	Latin	Italian	Malagasy
gain from word order manuscript 0.40	-0.00	0.01	0.01	2.94	0.39	0.00	1.71	2.92	3.41
previous word helps (held out) manuscript -0.12	-0.31	-0.11	-0.38	-0.14	-0.06	-0.27	0.83	1.65	1.62
gain from glyph context manuscript 0.22	0.19	0.13	0.49	-0.03	0.12	0.15	1.19	1.11	0.91
spelling–use correlation manuscript 0.15	0.04	-0.01	-0.00	0.01	-0.03	-0.01	0.03	0.03	0.09
near-twins among neighbours manuscript 0.15	0.05	0.06	0.01	0.13	0.07	0.05	0.03	0.07	0.03
repeated 4-word phrases manuscript 1.00	0.00	0.00	0.00	1.45	9.65	0.00	257	904	1694
near-twin within 10 words manuscript 0.26	0.19	0.03	0.00	0.36	0.22	0.12	0.07	0.18	0.13
adjacent-page similarity (z) manuscript 16.24	1.94	2.03	1.69	1.13	2.03	1.89	11.83	15.99	10.20
word-edge dependency (bits) manuscript 0.21	0.01	0.22	0.01	0.01	0.01	0.01	0.04	0.11	0.29
line-initial effect manuscript 0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Source: 20 simulated corpora per procedure, parameters fitted on h2, word length, TTR and hapax share; languages: Christodoulopoulos & Steedman Bible samples. Values in data/analysis/framework/out/generators_scores.csv.

Figure 25.2 Six procedures and three languages against ten held-out statistics of the manuscript.

26

Fragments

The searches of chapter 23 all assumed that the units of the cipher were known in advance: a glyph, a slot filler, a merged pair, a whole word. Every surviving cipher hypothesis says they are not. In a verbose cipher the written word is a string of glyphs that may be one plaintext unit, or two, or part of one, and the cut between units is nowhere marked. A solver that is to find such a key has to learn where the cuts fall at the same time as what the pieces mean. This chapter builds that solver, finds out on planted keys what it can and cannot recover, and then turns it on the manuscript.

THE SOLVER

The model is simple to state. A lexicon of candidate segments is drawn from the text: every string of one to eight glyph units that occurs at least fifteen times, which in the manuscript's paragraphs comes to about a thousand. Each written word is a concatenation of segments from the lexicon; each segment encodes one plaintext letter; many segments may encode the same letter; the plaintext has no spaces. The search alternates two steps. With the segmentation fixed, the key is fitted by the annealing of chapter 23, with two changes learned from the controls below: a stronger penalty against collapsing the alphabet onto a few letters, and a final pass in which every segment is moved to its best letter given all the others. With the key fixed, every word is re-cut by a beam search for the segmentation that makes the decoded letters most Latin-like under the four-gram model, the context carried across words. Four rounds of this were run; the key is fitted on the training quires and everything is scored on the held-out

quires.

Two scores are reported, and the second matters more. The first is the gap between the four-gram score of the decoded held-out text and the score of the language's own text, as in chapter 23. The second is a scorer the search never saw: the share of the decoded text that can be covered by dictionary words of four letters or more, found by a dynamic programme over the space-less string. A key that makes Latin-like four-grams without making Latin words scores well on the first and not on the second.

WHAT THE CONTROLS ALLOW

Before the manuscript, Latin. We enciphered Latin with an affix-table key of the Naibbe kind: each letter has four prefixes and four suffixes drawn from the manuscript's own inventory of word-initial and word-final strings, and each plaintext letter is written as a prefix followed by a suffix, so that a written word is one letter and the plaintext spaces are gone. The solver was given the ciphertext and nothing else. It recovered 81 per cent of held-out words to their plaintext letters, against 13 per cent for the commonest letter alone, and the decoded text reads *dominietcanuitexoculiseiusvidensque*: Latin without its spaces. Vowel-less Latin under the same kind of key was recovered at 87 per cent. The dictionary coverage of the decoded Latin is 51 per cent, against 86 for the plaintext itself, which calibrates the second score: a correct key found by this solver covers about half the text with dictionary words.

Then a harder key: the same tables, but two plaintext letters per written word, so that each word is a prefix, a suffix, a prefix and a suffix, and the solver has to find the cut in the middle. It did not. Five per cent of held-out words were decoded correctly, against two by chance; the decoded text drifts toward Latin four-grams without ever becoming Latin. The reason is in the model: a segment encodes one letter, and a word built as prefix and suffix of one letter decodes, under the right cut, to that letter twice. The solver can learn a segmentation but not the rule that a suffix encodes nothing, and without a null it cannot represent the key. This is the method's limit, and it is stated here so that the manuscript result can be read correctly: the search that follows can find a key in which each written word is one plaintext unit, cut however the text cuts it, and cannot find one in which written words concatenate two or more.

THE MANUSCRIPT

Table 26.1 gives the runs: the paragraphs against Latin, vowel-less Latin and Italian, and the ring text of the diagram pages, which chapter 24 showed to be a different register, separately against Latin and vowel-less Latin.

For Latin, the solver settles on 927 segments and cuts the paragraphs into 1.56 segments per word; the commonest segments are **y**, **ol**, **o**, **ch**, **l**, **dy**, **qo**, **or**, **r**, **daiin**, **ar**, **aiin**, **edy**, **chedy** and **shedy**: the suffixes and prefixes of the slot grammar, which is what a Latin four-gram model finds it easiest to map onto letters. The gap on held-out sheets is 0.78 nats, better than the 1.53 that permuted keys reach when they are allowed the same re-segmentation, and far from the control's 0.41. The decoded text is *aeietisciie-*

Table 26.1 The fragmenting-cipher solver. Controls: Latin enciphered with a planted affix-table key (four prefixes and four suffixes per letter); accuracy is the share of held-out written words decoded to their plaintext letters. Manuscript: gap is the language's own four-gram score minus that of the decoded held-out text, in nats per character; permuted is the same gap for the found key with its letters reassigned among the segments at random, each permuted key allowed the same re-segmentation (mean of six). Coverage: share of the decoded held-out text covered by dictionary words of four letters or more, a scorer the search did not use; for the controls, the plaintext's own coverage is given in brackets.

text	plaintext	units / word	segments	gap	permuted gap	coverage %	accuracy %
control	Latin	1	–	0.42	–	50 (86)	81 (chance 13)
control	Latin	2	–	0.64	–	38 (86)	5 (chance 2)
control	Latin, no vowels	1	–	0.13	–	79 (79)	87 (chance 15)
paragraphs	Latin	1.56	927	0.78	1.53	33 (permuted 32)	–
paragraphs	Latin, no vowels	1.47	722	0.35	0.38	72 (permuted 66)	–
paragraphs	Italian	1.57	953	0.91	1.49	32 (permuted 31)	–
rings	Latin	2.06	159	1.04	1.78	32 (permuted 26)	–
rings	Latin, no vowels	2.03	133	0.29	0.34	72 (permuted 69)	–

limeaeeiantiasutanitatiinusaegeduueetst: Latin letters in Latin proportions, with Latin-looking runs, *anitatiinus*, *tenuentame*, and not one Latin word. The dictionary coverage says so without needing to be read: 33 per cent, against 32 for the permuted keys and 51 for the control. Italian gives the same picture, 0.91 nats against 1.49 permuted, coverage 33 against 31. Vowel-less Latin, the nearest miss of chapter 23, is here indistinguishable from its permutations, 0.35 against 0.38 nats, 72 per cent coverage against 66, because a dictionary of consonant skeletons covers almost any string of **t**, **s**, **n**, **m** and **r**. The rings, cut at two segments per word, do no better: 1.04 nats against 1.78, coverage 32 against 26.

The solver has therefore found, in the manuscript, the best way to turn the slot grammar into Latin letter frequencies, and nothing beyond it. The score it reaches above its permutations is the score of a good letter-frequency match, not of a reading, and the scorer the search did not use cannot tell its output from a permuted key. Under the one hypothesis this solver can test, that each written word is one plaintext unit of Latin, Italian or vowel-less Latin, there is no key.

WHAT REMAINS OPEN

The hypothesis the solver cannot test is the one that chapter 25 left as the most likely: units of several sizes under one key, written two or three to a word, some segments encoding nothing. Nothing in this chapter bears on it. The next solver needs a null, so that a suffix can be a flourish; it needs segments that can encode a syllable or a word as well as a letter; and it needs to be validated on planted keys of that shape before it is believed on the manuscript, since this one passed its first control and failed its second. We have been explicit about the failure because it marks the boundary of what has been excluded: everything up to one unit per word, and nothing beyond.

A last observation belongs to the record. In every run the solver's segmentation converged on the same pieces, the prefixes and suffixes that the slot grammar of chapter 6 describes and that the hidden Markov model of chapter 22 rediscovered as onsets and codas. Three independent methods, one of them looking for Latin, now agree on where the cuts in

a Voynich word fall. If there is a cipher, those are its units, and any proposed key that does not assign them a role has to say why.

27

Lists and Alphabets

Two objections could be made to everything in this part, and this chapter answers them with data. The first is that the manuscript's weak word order might be a property of its genre rather than of its encoding: an inventory, a glossary or a star table carries little order between its items, and the pictures say this is a book of plants, stars and remedies. The second is that every statistic in this book was computed on one transliteration, and a transliteration is a theory of what the glyphs are. If the choice of units drives the results, the results are about the transliteration.

ARE LISTS LIKE THIS?

We took six Latin texts of the genres the pictures suggest: Apicius's recipes; Cato's *De agri cultura*, a farmer's manual of instructions and recipes; the first book of Pliny's *Natural History*, which is its table of contents, a list of 2,500 entries; books 12 and 17 of Isidore's *Etymologies*, the glossary-like chapters on animals and plants; and Hyginus's *Astronomica*, a catalogue of the constellations. Against them we set two prose texts, Pliny's second book and a sample of the Vulgate, and the manuscript itself, sampled as runs of consecutive pages at each text's length so that the comparison is not confounded by size.

Figure 27.1 shows the two order statistics of chapter 23 for each. The lists are not where the manuscript is. They are, on the whole, *more* ordered than prose: Apicius gains 2.3 bits per word from its word order, the Pliny index 2.5, Cato 1.4, Hyginus 1.2, against 0.6 for Pliny's prose; and they repeat phrases more, not less, with 60 to 310 repeated four-word sequences against 18 in the prose and

none in the manuscript. A recipe collection is formulaic by nature: *teres piper, ligusticum, origanum* recurs, and the index repeats *de* and *natura* in every entry. The glossaries of Isidore are the least ordered of the Latin texts, at 0.6 to 0.8 bits, and even they gain from the previous word on held-out pages, where the manuscript loses. At the manuscript's matched lengths its order gain is 0.27 to 0.37 and its held-out bigram gain -0.13 to -0.16; no Latin genre sampled approaches either figure, and none combines weak order with the absence of phrase repetition. Genre does not explain the order; the encoding does.

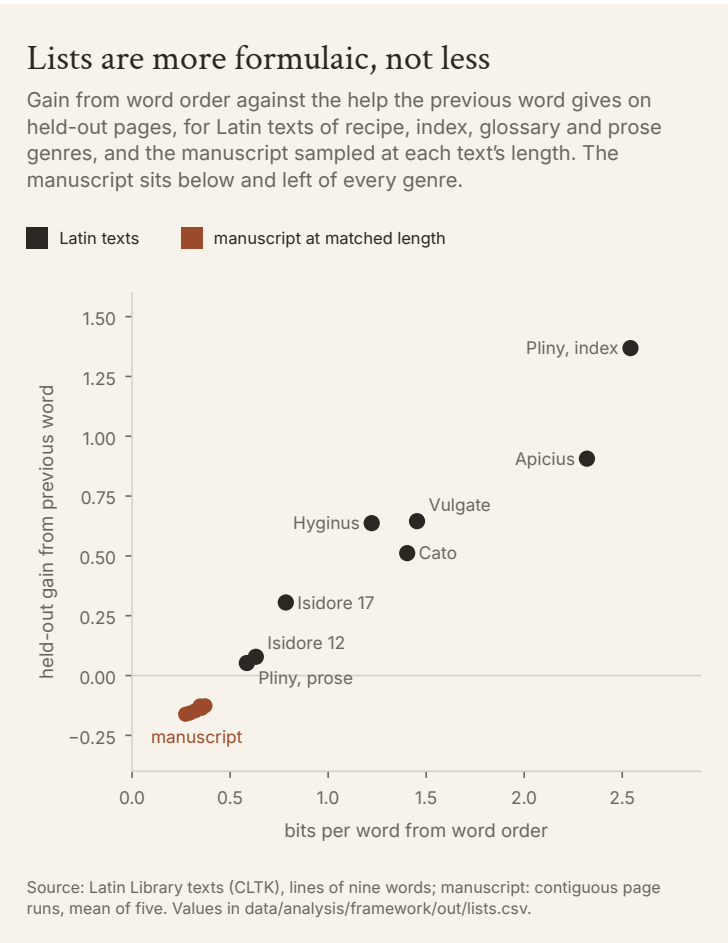


Figure 27.1 Order statistics of Latin list genres against the manuscript at matched length.

DOES THE ALPHABET MATTER?

The second objection is the more serious one, and the literature has long been aware of it: Reddy and Knight, Hauer and Kondrak, and Bowern and Lindemann all note that results shift with the transliteration, and chapter 5 explained why the glyph inventory is uncertain (Bowern and Lindemann, 2021; Hauer and Kondrak, 2016; Reddy and Knight, 2011). Our tests were run on the Zandbergen–Landini file in EVA, with 29 glyph units. The obvious check is to run them again on a file that makes different choices, and the most different available is Glen Claston's v101 alphabet, in which each distinct shape is one character and the inventory has 69 symbols; Zandbergen distributes it in the same format. Every test of this part that does not depend on EVA-specific units was re-run on it with one switch and no other change.

The word-level results do not move (figure 27.2). The share of distinct words is 0.23 against 0.22; the hapax share 0.70 against 0.71; the Zipf slope -1.01 against -1.04; the gain from word order places the manuscript 2.8 standard deviations off the entropy law against 2.9; the previous word makes the next harder to predict on held-out pages in both; the coupling of spelling to use is 0.16 against 0.15, and the share of near-twins among nearest neighbours is higher, 24 against 15 per cent; the ring text has 1.8 per cent **q**-initial words against 14.5 in paragraphs, as before; there are three repeated four-word sequences instead of none; the periodicity, isomorph and canonical-order tests give the same verdicts; and the nearest-language families at each level are the

same.

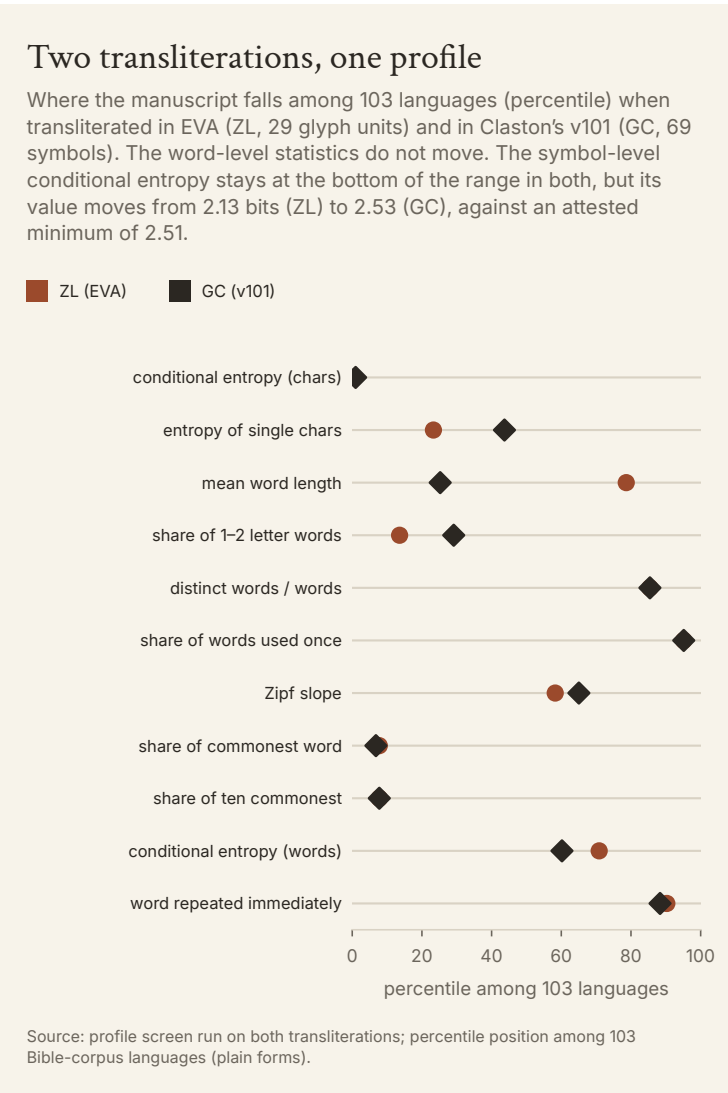


Figure 27.2 The profile under two transliterations.

The symbol-level results do move, and the book should say so. Under EVA the conditional entropy of the characters is 2.13 bits, far below the lowest of 103 languages at 2.51; under vIOI it is 2.53, at the edge of the attested range. The difference is what one expects when a transliteration splits shapes that EVA merges: more symbols, each rarer, each less predictable from the last. The anomaly that Bowern and Lindemann found robust across 250 languages is robust in sign but not in size, and the test of chapter 23 that excluded single-glyph nulls on the ground that no deletion lifts the entropy into the attested range does not hold under vIOI, where deleting the two commonest symbols lifts it to 2.75. We leave the EVA results standing in the earlier chapters, with this chapter as their qualification: what the manuscript's words do is the same under both alphabets; what its symbols do depends on what one counts as a symbol.

Part VII

The Chancery Key

28

Chancery Keys

CHAPTER 25 ended on a description of a key nobody had built. A procedure that fits all the evidence would write the same plaintext item the same way often enough for the vocabulary to cluster by subject, and vary it often enough that phrases never recur and word order carries nothing. The keys that the chanceries of Milan and Mantua were issuing in the decade the parchment was made come close to that description: a table of letters, syllables and whole words under one key, each with a few signs. This part builds that key, writes the manuscript with it, and looks for it in the manuscript with methods borrowed from the people who break such keys for a living and with some made for the purpose.

WHAT THE KEYS WERE LIKE

Most of what is known about fifteenth-century Italian keys comes from the ledgers in which secretaries copied them, and from the few ciphered letters that have been read against them. Aloys Meister printed the papal keys from the fourteenth century on (Meister, 1906). In the 1360s Archbishop Petrus of Naples replaced only the vowels, and scattered meaningless signs among the others "at will". Gabriele de Lavinde's keys of 1379 give every letter one sign, list four to six nulls under the heading *nichil habent importare*, and in the largest key add forty or fifty names and titles, each written as a short group of letters. The Tranchedino ledger in Vienna, copied around 1475 from keys of the previous decades, holds 287 complete keys. In the ones Nick Pelling has described, vowels get three signs, consonants two,

there are signs for doubled letters and for the common abbreviations, a few nulls, and a nomenclator of more than eighty entries; the Mantuan key of 1450 adds a series of syllables, *ab*, *ac*, *ad*, each with its own sign, and a small code for words such as *come*, *quando* and *non* (Pelling, 2016).

The keys are one half of the evidence; the letters written with them are the other. Giorgio Costamagna read a Milanese letter of 24 December 1453 against its key (Costamagna, 1977): more than forty signs in use, homophones on the vowels only, nulls "practically not used", a few nomenclator signs, and no word divisions at all. Cicco Simonetta's rules for breaking ciphers, written at Pavia in 1474, attack the endings of words and so assume that the divisions were visible (Simonetta, 1474); Meister found that papal keys began to forbid word spacing only in the 1540s. Both practices were in use. Across 700 European keys of the fifteenth to eighteenth centuries, Beáta Megyesi and her colleagues found nulls in 96 per cent of those from the fifteenth century, nearly always with two or more signs each (Megyesi et al., 2021); whole-word codes in that century were mostly names, places and titles, with function words in a minority of keys (Megyesi et al., 2022). Syllable signs were not yet common.

Three things follow for a simulator. The parameters are bounded: one to four signs per vowel, one or two per consonant, no syllables or a series of up to about eighty, a nomenclator of ten to a hundred entries, four to six kinds of null at rates from nothing to one sign in ten. The spacing practice is open, so it has to be simulated three ways: one written word for each plaintext word; plaintext words broken across written words; and a continuous stream cut into words by the scribe. And the signs themselves were graphic: invented shapes, letters, numerals, often with a dot or a stroke. A key of this kind written in Voynichese would need its signs to be strings of the manuscript's glyphs, because the manuscript has some twenty-five distinct glyphs where a Tranchedino key has fifty to a hundred and fifty signs. René Zandbergen raised exactly this objection in 2011. It does not exclude the hypothesis; it says what form the hypothesis has to take.

THE HYPOTHESIS, AS REGISTERED

Before any of the tests in this part were run, the hypothesis and four predictions were added to the registry and the file frozen (its SHA-256 digest is printed in Appendix E). The hypothesis is that the text is Latin, or an Italian vernacular, enciphered with an Italian chancery key of this mixed type, under any of the three spacing practices. The predictions:

1. **Fingerprint.** A simulator of the family, fitted on the training quires, reproduces both the weak word order (order gain within 0.25 of the manuscript's 0.42, held-out bigram gain at or below zero) and the topical locality (z within 2 of 16), which no procedure of chapter 25 reproduced together, and the coupling of spelling to use (at least 0.10).
2. **Collapse.** Merging spelling variants that behave alike in context gives back word order and repeated phrases faster than merging the same number of random variants, and faster than the same procedure does to Latin.

3. **Choice.** The way successive occurrences of a word's variants follow each other within a page, cycling, random or sticky, is measurable and the same for each scribe.
4. **Solver.** A solver that recovers planted keys of the family from Latin, at half or more of held-out words or seventy per cent of the plaintext's dictionary coverage, decodes the manuscript's held-out quires above its own permutation null by more than the control's margin. If the solver fails its controls at some rung of a ladder of increasingly complete keys, the hypothesis is reported as untested beyond that rung.

A KEY IN THE MANUSCRIPT'S SCRIPT

The simulator takes Latin plaintext of the manuscript's genres, herbal, astronomical, balneological, pharmaceutical and recipe text, section by section, so that a page of herbal cipher encodes herbal Latin, and enciphers it with a key built from the manuscript's own material. Each sign is a string of glyphs that the manuscript actually writes in one position of a word: a sign used at the start of a plaintext word is drawn from the strings that begin Voynich words (**qok**, **ch**, **o**), one used in the middle from the strings that fill the middle (**ee**, **ol**, **a**), one used at the end from the endings (**y**, **aiin**, **dy**). This gives the simulated text the manuscript's slot structure inside the word, which a key of arbitrary signs would not have, and it is the most favourable form of the hypothesis we could construct. Vowels get k_v signs and consonants k_c; the commonest letter pairs, and optionally the commonest triples, get one sign each; the commonest words get a nomenclator sign shaped like a whole Voynich word; nulls are inserted at a fixed rate. The encipherer chooses among homophones at random, by cycling or by sticking to the last one used.

THE FAMILY, MAPPED

Before testing the single key that fits the manuscript best, we wanted to know whether any key of the family could produce it. A grid of 216 keys covers the documented range: two, three or six signs per vowel with one, two or four per consonant; sixty or three hundred syllable signs, with or without two hundred triples; no nomenclator, eighty entries or three hundred; no nulls or one sign in ten; and each of the three spacing practices. Every key wrote the whole manuscript once, into its own pages, and every result was scored on the statistics of chapter 25 (figure 28.1).

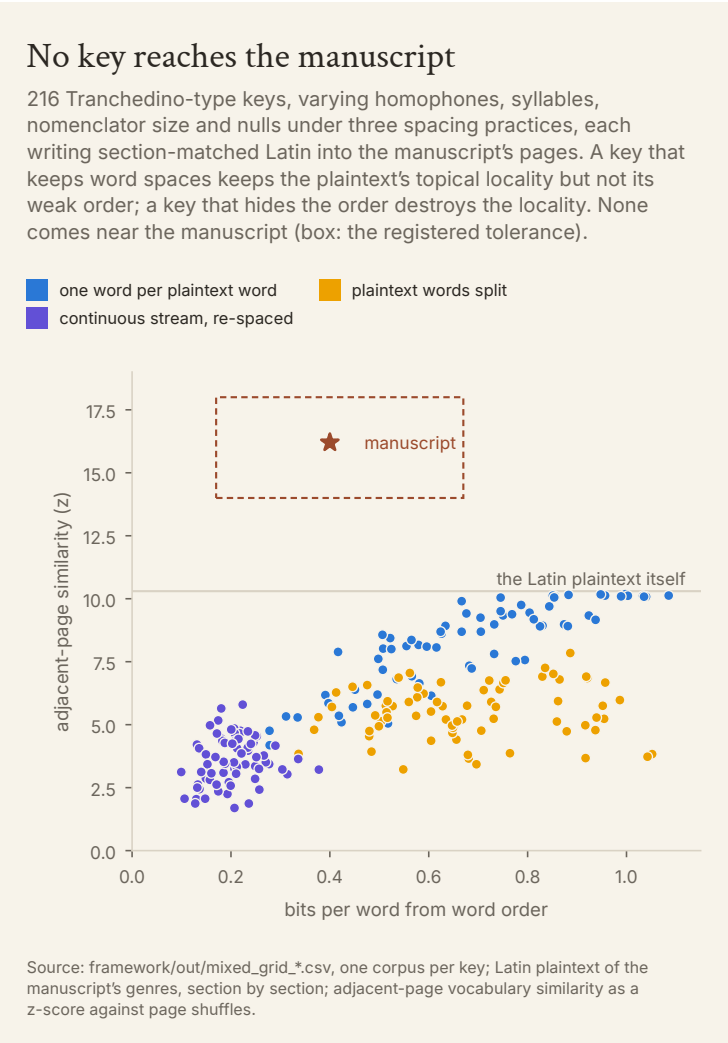


Figure 28.1 Every key of the grid, on the two statistics that pulled against each other in chapter 25.

The spacing practice decides almost everything. A key that keeps one written word for each plaintext word keeps the plaintext's topical locality, z up to 10.2 against 10.3 for the Latin it enciphered, and keeps its word order with it: the order gain runs from 0.2 to 1.1 bits, and in three keys of four the previous word still helps predict the next on held-out pages. Its words are also far too long, 8.5 to 13 glyphs where the manuscript has 4.6, because a Latin word of six letters becomes six signs of one to three glyphs each. Splitting plaintext words across written words brings the length down to the manuscript's, and gives the coupling of spelling to use that no earlier procedure produced (0.11 in the median key, 0.21 at most, against the manuscript's 0.15), because the pieces of one plaintext word share their signs. But in all seventy-two split keys the previous word helps predict the next, since the next word is usually the rest of the same Latin word. A continuous stream re-spaced by the scribe hides the order completely, as the manuscript does, and produces the dependency across word edges (0.12 to 0.26 bits, against 0.21), but the locality falls to between 1.7 and 5.8. In no key of the 216 is the locality within two of the manuscript's 16, and in none is the line-initial effect above 0.002, where the manuscript's is 0.198.

Two cautions belong with this. The locality a key can keep is bounded by the plaintext's, and our Latin of the manuscript's genres reaches 10.3; a plaintext more strictly organised by subject, a herbal in which each page names one plant and its virtues, could give a key with word spacing a higher figure. That would not rescue word spacing, which fails on length and order regardless. And the grid is a map, not a test: each key was run once. The registered test takes, for each spacing practice, the key nearest the manuscript on the four fitted statistics of the training quires and writes the manuscript twenty times with it.

THE REGISTERED TEST

For word spacing the nearest key has two signs per vowel and one per consonant, sixty syllables, two

hundred triples and a nomenclator of three hundred words; it is still far from the manuscript on what it was fitted to (words of 8.6 glyphs, 41 per cent of them distinct). For split words the nearest key has three signs per vowel, two per consonant and three hundred syllables, and it fits well: words of 4.9 glyphs, 23 per cent distinct. For the stream, two signs per vowel, one per consonant and sixty syllables; words of 4.7 glyphs. Each was written twenty times. Table 28.1 gives the twelve held-out statistics.

Table 28.1 The three fitted keys, twenty corpora each, against the manuscript on held-out statistics. Mean (and the manuscript's distance in standard deviations of the simulations).

	manuscript	word	split	stream
gain from word order	0.40	1.20 (−5)	0.92 (−12)	0.39 (0)
previous word helps	−0.12	0.43 (−4)	0.57 (−17)	0.07 (−6)
glyph context gain	0.22	0.65 (−19)	0.60 (−20)	0.48 (−10)
spelling–use correlation	0.15	0.00 (18)	0.16 (0)	0.05 (6)
near-twin neighbours	0.15	0.00 (77)	0.03 (20)	0.05 (17)
repeated 3-grams	145	630 (−3)	294 (−3)	79 (4)
repeated 4-grams	1	297 (−3)	36 (−3)	6 (−2)
near-twin in 10 words	0.26	0.01 (100)	0.08 (34)	0.12 (14)
the same, shuffled	0.23	0.01 (89)	0.08 (29)	0.12 (11)
page locality (z)	16.2	10.8 (5)	4.7 (14)	3.4 (20)
word-edge dependency	0.21	0.01 (145)	0.20 (0)	0.22 (−0)
line-initial effect	0.198	0.001 (517)	0.001 (592)	0.002 (306)

The first prediction fails for all three practices. The word-spaced key fails all twelve statistics. The split key reproduces the coupling of spelling to use and the dependency across word edges, two things no procedure of chapter 25 did together, but its word order is worth more than twice the manuscript's and its locality is a third of it. The stream reproduces the weak word order and the edge dependency and nothing else; its pages are barely more alike than random ones. Every key fails the gain from glyph context, the near-twin rates and the line-initial effect by wide margins. The best any spacing practice does is three statistics of twelve, which is one better than the six procedures of chapter 25, and the trade-off between order and locality that ended that chapter is still there, now mapped across the whole documented range of the keys that were supposed to escape it.

29

Looking for the Key

A simulator can show that a key of some kind would produce text like the manuscript's. It cannot show that the manuscript was produced by one. For that the key has to be found in the text: its homophones as groups of spellings that stand for the same thing, the encipherer's habits in choosing among them, and in the end the key itself, recovered well enough that held-out pages read as Latin. This chapter tries all three, first on keys we planted, then on the manuscript.

WHAT A SECRETARY OF 1474 WOULD HAVE SEEN

The oldest method is the one the keys were designed against. Cicco Simonetta's thirteen rules begin with the ends of words: find the signs that end words, and count them; Italian has five or fewer endings that matter, Latin more, and the endings are mostly vowels (Simonetta, 1474). Applied to the manuscript, the rule is unambiguous. Eight glyphs end more than one word in a hundred, and the commonest five, **y**, **n**, **l**, **r** and **o**, end 91 per cent of all words; **y** alone ends 41. In the Italian Bible eleven letters pass the same threshold and the commonest five cover 83 per cent; in the Latin, twelve and 71. Simonetta would have called this a vernacular. The point for a key of his time is the direction: vowel homophones exist to spread the endings over more signs, and the manuscript's endings are more concentrated than plain Italian's. His next rule, that a sign always followed by the same sign is *q* with its *u*, also finds its case: **q** is followed by **o** in 97.6 per cent of its occurrences. Nick Pelling has long read **qo** as a cipher pair of the kind used in Milan and Modena in those

decades (Pelling, 2006), and Simonetta's rule points the same way.

MERGING THE VARIANTS

If **chol** and **chor** are two signs for one plaintext word, writing both the same way should give back some of the plaintext's structure: phrases that recur, an order between words that carries information. That is the idea of the collapse test. The pre-registered form of it merged whole words joined by a one-glyph difference and turned out to be degenerate: the merges chained, and three hundred of them put half of all tokens into a single family, which makes any text repetitive. We record that and the revision, which works at the level a homophonic key works at. Every pair of words that differ in exactly one glyph unit, **chol** and **chor**, is a minimal pair of the alternation **l**/**r**. An alternation is scored by how alike the contexts of its minimal pairs are, and the k highest-scoring alternations are merged; the null merges k alternations drawn at random. Merging anything shrinks the vocabulary, and a smaller vocabulary repeats more, so the comparison is made at equal numbers of distinct words, interpolating the random merges to the vocabulary the context merges leave.

At twenty alternations the manuscript's vocabulary falls from 7,473 words to 2,162, and repeated four-word sequences rise from one to 171; random merges that leave 2,260 words give 56. The context-chosen merges do produce phrases, three times as many as chance at the same vocabulary. But two things weaken this. Latin, put through the same procedure, also gains (294 repeated sequences against about 250 expected), because merging words that share neighbours makes the neighbours repeat; part of the effect is built into the measure. And the word order does not come back: the order gain after the merge is 0.27 bits per word, lower than before, and no higher than random merges give. On the simulated keys the test has almost no power. A key whose homophones are whole strings of glyphs, **qok** for one sign and **ch** for another, makes few minimal pairs that differ in one glyph: the key with word spacing gave three eligible alternations in some 34,000 words. The key with split words gave 144, but merging the best twenty of them brought its order gain to 0.34, where the true plaintext words give 0.83. So the second prediction fails on the manuscript, which does not regain order, and the instrument would not have seen a key of this family if one were there.

HOW THE VARIANTS ARE CHOSEN

Fourteen *e*s in the Zodiac's 408-symbol cipher are written as two identical runs of seven homophones: the killer cycled through his signs in order (Oranchak et al., 2024). A Milanese secretary of 1453 never wrote the same sign twice in a row (Costamagna, 1977). People choosing homophones by hand leave traces of that kind, and the third prediction was that the manuscript's choices would show one. Within each page, we followed each family of variants through its occurrences and counted how often a variant repeats the previous one, against the rate when the page's variants are put in random order.

The manuscript's ratio is 1.04: variants repeat slightly more than chance, and by about the same

amount for every scribe (1.02 to 1.08 for the three with enough text). But the measure failed its calibration. Simulated keys with random, cycling and sticky choice gave 1.01, 0.995 and 1.015 on their true families and 1.06 on the families the merge procedure recovers, whatever the choice. A written word is several signs, each chosen separately, so cycling at the level of the sign is diluted to nothing at the level of the word, and the families recovered from the text add their own bias. The prediction cannot be tested with this instrument; the manuscript's figure is reported, not interpreted.

THE LINE KNOWS

The choice of a homophone has one property that does not depend on what the encipherer's habits were. It is made when the word is enciphered, and the line break falls afterwards, where the page runs out. So within a minimal pair, which of **chol** and **chor** is written should not depend on where in its line the word stands. In a cipher of this family that is close to a law: it can fail only if the encipherer chose signs by the layout, or if the scribe changed them while copying.

We tested it within frames. For each of the ten commonest alternations, the words of each minimal pair were taken together and the variant written was compared with the word's position in its line, first, inside or last, by the mutual information between the two, summed over frames, against five hundred shuffles of the variants within page and frame. The first and last lines of paragraphs were left out, since their beginnings and ends are special for other reasons. The controls are Latin laid into the same lines, the three simulated keys, and, for the objection that a plaintext organised in lines might do it, the Vulgate laid out one verse to a line and enciphered with the same key, so that the first word of a line is the first word of a verse.

Figure 29.1 gives the result. In none of the fifty-three alternations tested on the keys and the Latin does the line position matter at three standard deviations, apart from one class of the continuous-stream key at 3.4; the verse-per-line cipher gives none above 1.4. In the manuscript six alternations of ten depend on it, at z = 3.9 to 9.2. The directions are not small: among words that differ only in **d** and **sh**, the variant with **d** is written 62 per cent of the time inside a line and 93 per cent of the time at its end; among those that differ in **l** and **q**, **l** is chosen 12 per cent of the time inside and 44 per cent at the end; **cth** against **sh** goes from 23 to 64. The effect holds in both of Currier's languages (four alternations of ten in each) and in the scribes' own pages, most strongly in scribe 1's.

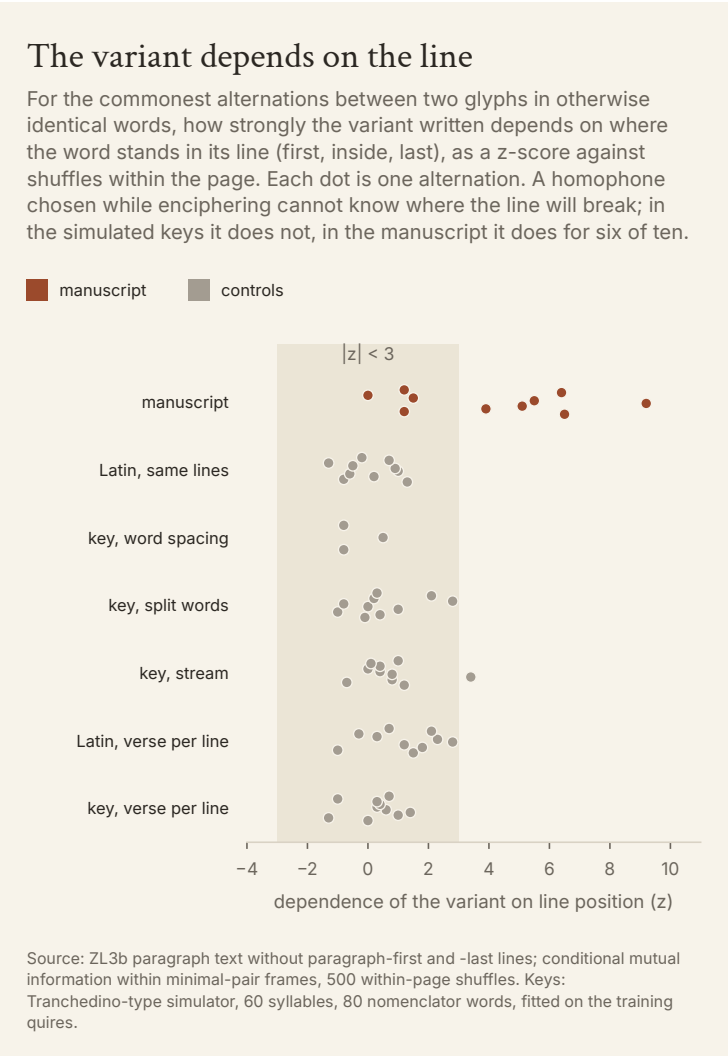


Figure 29.1 Dependence of the variant on line position within minimal pairs, for the manuscript and the controls.

This is the line-initial and line-final behaviour that chapter 7 described and that every procedure in chapter 25 failed to reproduce, now measured where a cipher has to be blind to it. It does not by itself exclude a key, because a scribe copying from a cipher exemplar could have adjusted words to the line, as copyists of Latin adjusted abbreviations. But it means that if there is a key, the text we have is not its direct output: something between the key and the page chose spellings by layout, and did it differently for different spellings. Any solver run on the manuscript as written is run on a text that has been altered at the ends of its lines.

A SOLVER FOR MIXED KEYS

Chapter 26's solver assumed that each segment of a written word stands for one letter, and it failed when segments stood for more. A key of this family needs a solver in which a segment can stand for a letter, a syllable, a whole word, a word boundary, or nothing. The first two objectives we tried, both scoring the decoded text letter by letter, failed on planted keys in the same way: the decoded text could be made longer or shorter by choosing long or short units, and the solver learned to write strings of common Latin fragments, *et in eis et in*, scoring well and meaning nothing, with between none and sixteen per cent of the planted key right.

The method that works is George Lasry's for the syllabic ciphers of the Italian chanceries (Lasry, 2024). The plaintext is treated as a sequence of elements, each of which is a letter, a syllable, a nomenclator word, a space or a null, and each cipher sign is assigned exactly one element. The language model is a model of element sequences, trained on a reference text decomposed into elements by the same scheme the key is assumed to use, with nulls sprinkled in at the expected rate so that the model expects them. Because every sign emits one element, the length of the decoded text is fixed by the ciphertext and cannot be bought. The search is simulated annealing over the assignment, with moves that reassign one sign or swap two, a penalty on element frequencies

that stray from the reference (the same role as the entropy term in AZdecrypt and as Lasry's concentration denominator), and a final pass that tries every element for every sign. When the segmentation of written words into signs is not known, it alternates with a beam search that re-cuts each written word into signs to maximise the score under the current key.

The ladder of controls plants keys of increasing completeness in 22,000 words of Latin, fits on the first three quarters and scores on the rest, with the language model trained on a different part of the Bible. The first run of the ladder had a fault in the simulator, not the solver: a glyph string could be a sign for one letter at the start of a word and for another at the end, so that the planted key was ambiguous, and the solver, which treats each string as one sign, stalled at the third rung. A diagnostic showed that the true key scored worse than the solver's own under the solver's objective, which is the signature of a model that cannot express the key. With every sign given one reading, as in the historical keys, the true key scores better and the solver finds it. The run with the fault is kept in the log; the results below are from the corrected keys.

Figure 29.2 shows the ladder. With the signs known, the solver recovers each part of the key in turn. Letters alone, three signs for each vowel and two for each consonant, each sign a string of the manuscript's glyphs: 98 per cent of sign uses assigned correctly, 89 per cent of held-out words read exactly. With four kinds of null at one sign in twenty: 97 and 86. With sixty syllable signs: 95 and 81. With a nomenclator of the eighty commonest words: 86 and 53, just above the pass mark. That is a full key of the Tranchedino type, with 389 signs, recovered from 17,000 words of ciphertext without a crib. That is in the range Lasry reports for real Italian ciphers (58 per cent of the key and 75 per cent of the text for Ottone del Carretto's of 1457), though ours is synthetic and the language model is of the same Latin. Where the plaintext of the fourth rung reads *ob rem adfligitur cor tuum numquid non ego melior sum*, the decoded text reads *os rem adfligitur cor tuum numuis in in melio sum*: the nomenclator words are where it goes wrong.

The fifth rung takes away the one thing a cryptanalyst of a real chancery cipher always has: knowing where one sign ends and the next begins. In a key of invented shapes that is given by the shapes. In a key written in Voynichese it is not, because a sign is a string of glyphs and a written word is several signs run together: **qokeedy** could be **qok-ee-dy** or **qo-kee-d-y**. The solver now has to find the signs too. Chapter 26's lexicon, built to find one unit per written word, finds the wrong ones (boundary F_1 0.31). A unigram segmentation of the kind Aldarrab and May used on unspaced numerical ciphers does much better (Aldarrab and May, 2022): it places 88 per cent of the boundaries correctly, whether or not plaintext spaces were kept. But only 76 to 83 per cent of sign tokens come out exactly right, and the solver, given that segmentation and six rounds of re-cutting it under the current key, recovers a quarter of the key and reads 9 per cent of held-out words. Without word spaces it recovers a fifth.

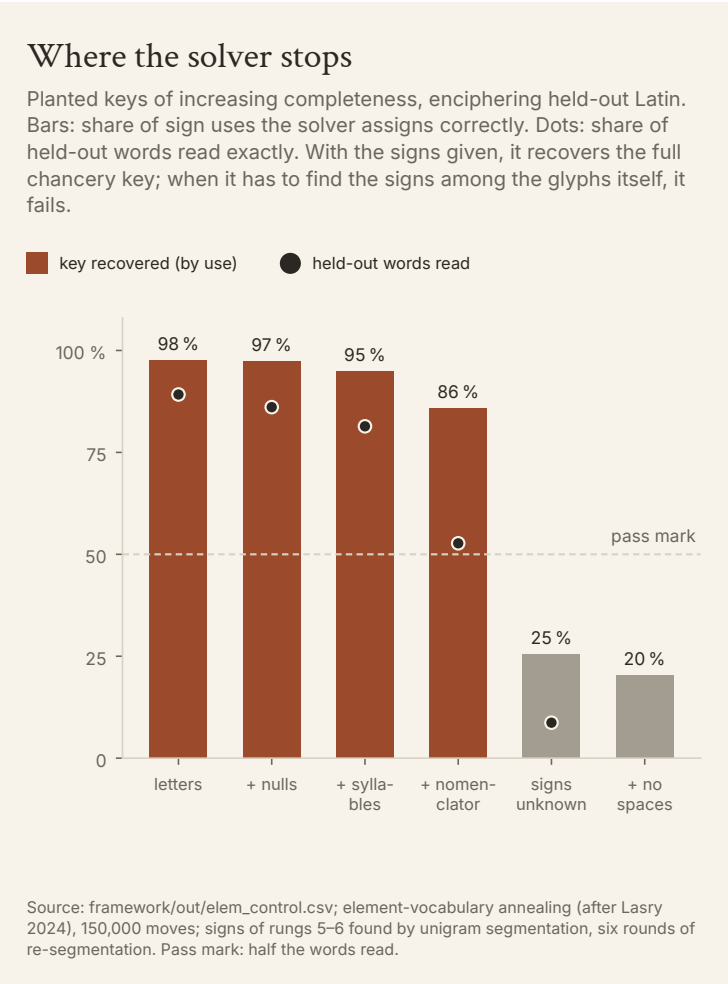


Figure 29.2 The planted-key ladder: share of the key and of held-out words recovered at each rung.

The failure is not one of search. Started from the best key the learned segmentation allows (each found unit given the reading of the true sign it most often coincides with), the joint procedure walks away from it: re-cutting the words under the key lowers the share of exactly-right sign tokens from 82 to 60 per cent within two rounds, and the key falls apart with it. Given freedom to choose both the signs and their readings, the model finds Latin-looking element sequences more easily in wrong cuts than in right ones. A solver of this design cannot recover a mixed key whose signs are strings of a twenty-odd-glyph script, even when we planted it, and by the rule registered beforehand, the hypothesis is untested beyond the fourth rung: it is testable only in a form, signs known, that the manuscript does not offer.

THE MANUSCRIPT UNDER THE SOLVER

The registered rule ends the test at the fourth rung, but the solver was run on the manuscript anyway, with the same unit discovery, six rounds and the same language models, because a failed control still leaves a question worth asking: does the manuscript look to the solver more like a cipher it cannot break, or more like text with no plaintext at all? The key was fitted on the training quires and the held-out quires decoded. The comparison is with the same text after every word type's glyphs have been put in a fixed random order, which keeps the vocabulary, the word frequencies, the word order and the glyph frequencies and destroys only the spelling; and, as texts that have the manuscript's word structure and certainly no plaintext, with the slot grammar and the self-copying generator of chapter 25, each run the same way.

Dictionary coverage, the scorer chapter 26 relied on, fails here even on the control: the planted cipher decodes to 33 per cent coverage and its permuted twin to 36. The language model's own score on the held-out text does separate them. For the planted cipher at the fifth rung, real spelling beats permuted spelling by 0.50 to 0.60 nats per element (two seeds), although only a quarter of the key is right. For the manuscript the margin is 0.18 to 0.27 against Latin

and 0.44 against Italian; for the slot grammar it is 0.00 and for self-copying 0.02. Without word spaces, the control's real text scores below its permuted twin and so does the manuscript's, and nothing can be read from either.

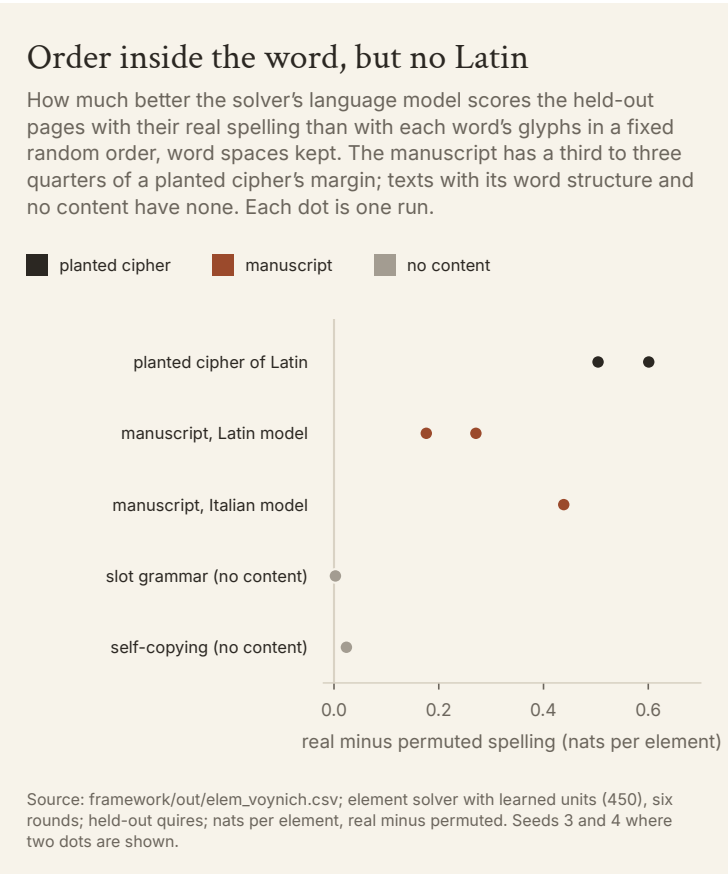


Figure 29.3 How much better the solver models real spelling than spelling with each word's glyphs shuffled, on held-out pages.

Figure 29.3 shows the margins together. The manuscript's spellings carry an order inside the word that the solver can use to make its output more Latin-like or Italian-like, about a third to three quarters of what a planted cipher of Latin gives at the same rung, and the two generators without content have none. That is what the fingerprints of chapter 25 showed by other means, where the manuscript gains more from glyph context (0.22 bits) than the slot grammar (0.15) or self-copying (none), measured now through a decipherment model. It is not a reading. The decoded held-out text begins *besai ae super aa a e et meea de ime a erit ti mute sum*, its coverage by Latin words is lower than its permuted twin's (27 against 29 per cent), and the larger margin is for Italian, which suggests that what the model finds is short vowel-rich sequences rather than a language.

THE VERDICT

The mixed key of the Italian chanceries was the strongest candidate the earlier parts left standing. Of its four registered predictions, the fingerprint fails for all three spacing practices across the documented range of the keys; the collapse fails on the manuscript, which does not regain its word order, and could not have seen a key of this kind on the simulated texts; the choice dynamics could not be measured; and the solver, which recovers a full chancery key when the signs are given, cannot recover one whose signs are strings of glyphs, so the hypothesis is untested at the rung that matters. The registry records it as not supported in its simulated forms and untested as a cipher with unknown sign boundaries.

Two results of this part reach beyond the hypothesis. A cipher written in an alphabet of twenty-odd glyphs, with signs of one to three glyphs and a few hundred of them, is not recoverable by the methods that read the real chancery letters, because the sign boundaries become part of the problem and the problem then has too many good-looking answers. That

is a limit on what can be learned from the text by decipherment of this kind, and it holds whatever the manuscript is. And the spelling depends on the line, within pairs of words that a cipher would treat alike: whatever produced the text, it was finished at the page.

30

Where This Leaves the Reading

Chapter 18 gave a reading of the manuscript that could be defended line by line, and chapter 19 a set of tests for anyone who claims more. Parts V to VII have added twenty-five registered hypotheses, tested against 103 languages, six kinds of surrogate, planted-key controls, six simulated procedures and 216 simulated chancery keys, and it is time to say what they change. The count of running-text words rendered as dictionary words of any language on held-out sheets, which chapter 20 set as the measure of progress, is still zero. What has changed is the shape of the problem.

WHAT IS NEW

Eleven results of these three parts were not, as far as we can find, in the literature before, and each can be checked from the files distributed with the book.

The sheet, not the binding. Pages on one leaf or one sheet share vocabulary at nine to ten standard deviations above chance; pages merely bound next to each other on different sheets do not, at one. The unit of composition is the sheet, and the present order of sheets carries no textual signal. (Chapter 22.)

Words that look alike behave alike. Among the 300 commonest words, spelling similarity and similarity of use correlate at 0.15, and 15 per cent of a word's nearest neighbours in use are its near-twins in spelling; no language in a sample of 21 exceeds 0.09 and 9 per cent, and no simulated procedure reaches it either. (Chapters 22, 25.)

No repeated phrases. Ninety-seven repeated three-word sequences and no repeated four-word sequence in 34,000 words, where the least formulaic

Latin tried has 128 and 16, and the recipe and index genres have hundreds. (Chapters 22, 27.)

Word order carries almost nothing. Knowing the order of the words saves 0.42 bits per word over the shuffled bag, against 1.34 for the lowest of 102 languages and a median of 2.4; on held-out sheets, knowing the previous word makes the next harder to predict, not easier. The result holds under a second transliteration and is not a property of list genres, which are more ordered than prose. (Chapters 23, 27.)

Weak order is a cipher's fingerprint too. A Naibbe-type affix-table cipher of Latin with large tables, fitted to the manuscript, reproduces its worthless word order exactly, because dozens of forms per letter hide the plaintext's order behind the table's randomness. Weak order excludes a cipher in which the written word is the plaintext word; it does not exclude a verbose cipher. (Chapter 25.)

But the words recur by subject. Adjacent pages are alike in vocabulary at $z = 16$, as languages are (10 to 20) and as no simulated procedure is (1 to 2). Whatever the words are, the same item tends to come out the same way often enough for the vocabulary to keep its topical shape. The larger a cipher's tables, the less this can happen; and the manuscript has it. (Chapters 22, 25.)

Rings are written like labels. Text running round the circles has the word-initial profile of labels, not of paragraphs, under both transliterations. (Chapters 24, 27.)

The hands are interchangeable. Three scribes writing the same section in the same language choose among variant spellings at the same rates ($z = 1.2$ against a page-shuffled null), while a change of subject moves the same statistics by nineteen standard deviations. Rote execution of a fixed procedure, or copying from one exemplar; not free composition. (Chapter 25.)

The symbol-level anomaly depends on the alphabet. The conditional entropy of the characters is 2.13 bits under EVA and 2.53 under v101, against an attested minimum of 2.51; the word-level results do not move. (Chapter 27.)

The chancery key does not escape the trade-off. A simulator of the Italian mixed keys of 1400–1475, letters with homophones, syllable signs, a nomenclator and nulls, its signs drawn from the manuscript's own glyph strings, mapped over 216 keys and three spacing practices: keys that keep word spaces keep the plaintext's locality and its word order; keys that hide the order destroy the locality. The fitted keys fail nine to twelve of twelve held-out statistics. (Chapter 28.)

The line chooses the spelling. Among words that differ in one glyph, which variant is written depends on the word's place in its line for six alternations of ten, at $z = 3.9$ to 9.2, with paragraph boundaries excluded; no simulated key and no Latin text, including the Vulgate set one verse to a line and enciphered, shows it. A homophone chosen while enciphering cannot know where the line will break, so the text as we have it was shaped at the line after any encoding. (Chapter 29.)

What is excluded, with controls. A one-to-one substitution of any of 103 languages; a homophonic substitution of Latin, Italian, German, French or Czech over four kinds of unit, with the frequency

penalty weak or strong; a word-as-letter cipher over the 300 commonest words; a fragmenting cipher in which each written word is one plaintext unit, with the segmentation learned, for Latin, vowel-less Latin and Italian; a periodic transposition; a position-keyed alphabet; a Tranchedino-type mixed key of Latin as simulated, under any of three spacing practices; the sorted anagram; a paint schedule by quire; and, as specified and fitted, each of six mechanical procedures taken singly. (Chapters 21, 23, 24, 25, 26, 28.)

WHAT THE RESULTS SAY TOGETHER

A text that is tightly ordered inside the word, organised by sheet and subject, nearly without order between one word and the next, written by five interchangeable hands. Each part of that has an explanation; the difficulty is that the explanations pull against each other.

The weak order and the interchangeable variants are what a verbose cipher with large tables produces, and chapter 25 showed one producing them. The topical locality and the coupling of spelling to use are what a text whose words keep their identity produces, and a large table destroys both. Part VII built the attested key that was supposed to reconcile the two, the Tranchedino type with letters, syllables and whole words under one key, and it does not: across the whole range of the keys the chanceries used, a key that keeps the locality keeps the word order, and a key that hides the order loses the locality. Its best member reproduces three of twelve held-out statistics. The key could not be looked for directly either. The solver recovers a full planted key of this kind when it is told where the signs begin and end, and fails when it has to find the signs among the glyphs, which is the situation the manuscript presents; run on the manuscript all the same, it finds more order inside the words than texts without content have, about a third of what a planted cipher shows, and no Latin.

The new observation of Part VII points somewhere else. Within pairs of words that differ by one glyph, the spelling chosen depends on the word's place in the line. A homophone chosen in the act of enciphering cannot depend on a line break that comes later, and Latin does not do it even when every line begins a verse. Whatever produced the text, the last hand to touch it chose spellings with the line in view: a copyist adjusting words to the space, as scribes of Latin adjusted their abbreviations, or a procedure executed directly on the page, line by line. Either way, any attempt at a key has to undo a layer of layout first, and every solver so far, ours included, has been run on the text with that layer still in it.

The alternative, that there is no plaintext and the text was generated by a procedure, has lost ground in this part. Self-copying reproduces the near-twin statistics and nothing else; the grille and the slot grammar reproduce almost nothing; and none of them, nor any procedure we can think of that does not carry content, produces pages that resemble their neighbours by subject at sixteen standard deviations. A generating procedure would have to be steered by the pictures, page by page, and a procedure steered by content is a way of writing content.

NEXT EXPERIMENTS

The list of the previous edition has been worked through as far as the mixed key; this one replaces it.

1. **The layout layer.** Model the line effect directly: for each alternation, the variant written as a function of line position, scribe and the glyphs around it; then a text with line-conditioned variants normalised to their inside-the-line form, and the battery and the solvers run again on that. 15–20 hours.
2. **A copying layer.** Every simulator, the mixed key included, passed through a model of a copyist who re-lineates and adjusts spellings at line ends, to see whether copying from an exemplar with different line breaks gives the manuscript's line effects and the locality together. 15 hours.
3. **A solver that can find the signs.** The fifth rung of the ladder is where the mixed key becomes untestable. Constrain the segmentation by the slot grammar, or decode over a lattice of all segmentations with a neural language model as Chu, Valenti and Knight did for dictionary codes, and validate on the same planted keys before any run on the manuscript. 30–40 hours.
4. **A more topical plaintext.** The locality a key can keep is bounded by its plaintext's; a herbal in which each page names one plant would show whether word-spaced keys can reach 16. 8 hours.
5. **Edge dependency decomposed**, **morphology induction**, **Ponzi's zodiac matrix as data** and **water across the gutter**, as in the previous list. 60–80 hours together.

THE READING, RESTATED

Nothing in these three parts has weakened the reading of chapter 18, and four things have strengthened it: the sheet as the unit of composition; the pictures and labels as the frame of a compendium organised by subject; the text as something whose words are built to a pattern and placed with little regard to one another; and the production as a procedure, carried out identically by every hand that worked on it. What a reading cannot yet say is what a word means. Part VI narrowed where the answer could be, to a key of the kind the chanceries of Milan and Mantua were using in the decade the parchment was made. Part VII built that key, measured it, and tried to find it. As simulated, it does not produce the manuscript; as a cipher whose signs are strings of this script, it could not be recovered from text even when we planted it; and the manuscript's spellings were chosen with the line in view, which a key used in the ordinary way does not do. The question of a key is still open, but it now has a precondition: the layer that the page added has to be taken off before anything underneath can be read.

A

Data and Methods

This appendix describes the data behind every number in the book and the programs that produced them, so that anyone can repeat the work. The programs are short, written in Python 3 with NumPy and Matplotlib, and are distributed with the book's source. Running the whole set, including the figures, takes under two minutes on an ordinary laptop.

SOURCE TEXT

ZL3b-n.txt. The Zandbergen–Landini transliteration of the manuscript in EVA, version 3b of 13 May 2025, in IVTFF format 2.0, distributed at `https://www.voynich.nu/data/`. It records 5,385 loci (lines of text, labels, circular and radial text) on 227 pages and panels, the rosettes fold-out being one page identified as `fRos`, with page variables giving each page's quire (`$Q`), position in quire (`$P`), illustration type (`$I`), Currier language (`$L`), Davis scribe (`$H`) and Currier hand (`$C`), and markers for the start (`<%>`) and end (`<$>`) of paragraphs.

IT2a-n.txt. Takeshi Takahashi's transliteration in the same format, used only for the agreement measure in chapter 5.

CD2a-n.txt. The Currier–D'Imperio transliteration, used only to confirm the correspondence between Currier's alphabet and EVA.

READING THE FORMAT

The reader (`ivtff.py`) applies these rules to each locus. In-line comments `<...>` are removed. High-ASCII codes `@nnn;` for rare glyphs are removed. Ligature braces `{}` are removed and their contents kept. Where the transcriber gives alternatives `[a:b]`, the first is taken. Uncertain spaces `(,)` are treated as

spaces. Lines continued with a trailing / are joined. Words containing ? (an unreadable character) are excluded from word counts; there are 140 in the running text.

"Running text" means loci of generic type P (paragraph text) only. Labels are loci of type L; circular and radial text are C and R. All counts in the book say which they use.

The first phase of the work used a reader that did not recognise the page identifier `fRos` and so omitted the rosettes fold-out: 37 paragraph words, 337 words in rings and 152 labels. The reader was corrected in the second phase and the statistics of Parts II–IV recomputed; no figure printed in those parts changed at its printed precision (the running text grew from 34,180 to 34,217 words). The tests of Part V include the page.

GLYPH UNITS

Statistics are given on the EVA characters and, where stated, on a "glyph-unit" alphabet in which the sequences `cth`, `ckh`, `cph`, `cfh`, `ch` and `sh` are counted as single symbols. No other merging is done; `iin` and `ee` are left as sequences.

REFERENCE TEXTS

Five samples of exactly 34,180 words were taken from public-domain texts: the Latin Vulgate and the Bible in English (King James), German and Italian from the parallel Bible corpus distributed by C. Christodoulopoulos and M. Steedman (`https://github.com/christos-c/bible-corpus`), and Books 2–5 of Pliny's *Natural History* from the Latin Library as distributed by the Classical Language Toolkit. Each sample is a contiguous stretch beginning one fifth of the way into the source (one twentieth for Pliny), lower-cased, with punctuation and digits removed. The samples are chosen for availability and size, not as candidates for the manuscript's language.

MEASURES

Entropy. $h_1 = -\sum_i p_i \log_2 p_i$ over single symbols; $h_2 = -\sum_{ij} p_{ij} \log_2 p_{j|i}$ over adjacent pairs; the word space is included as a symbol in both. *Index of coincidence:* $\sum_i n_i(n_i-1)/N(N-1)$, excluding the space. *Zipf slope:* least-squares slope of log frequency on log rank over ranks 1–1000. *Keyness:* signed log-likelihood G^2 for each word's counts in two groups, words with fewer than 20 occurrences combined excluded; the share of "significantly skewed" tokens is the share lying in words with $|G^2| > 10.83$ ($p < 0.001$). *Group similarity:* cosine between frequency vectors over the 400 commonest words. *B-score:* per page, the share of words containing `ed`, ending in `eey`, or beginning with `qok`, minus the share beginning with `cho` or `sho`, equal to `daiin`, or ending in `ol` or `or`. *Near-twins:* two words are near-twins if their edit distance is 1. *Substitution search:* hill-climbing over one-to-one assignments of the n commonest glyph units to the letters of the reference alphabet, scored by $\sum_{ij} C_{ij} \log P_{m(i)m(j)}$ where C is the Voynich bigram count matrix and P the reference bigram model with add-0.5 smoothing; 12 restarts of 6,000 proposals each, with the word space fixed.

FIGURES

All data figures were drawn by a small library (`vizkit`) that lays out each figure as a stack of measured text and plot regions at the final printed width, so that nothing is rescaled in the page, and that checks every rendered figure for overlapping or clipped text and for type below 6 points before it is written. The categorical palette was validated for colour-vision deficiency. The folio images are the Beinecke Library's IIIF images at 2,000 pixels wide, and the large fold-outs at full resolution.

FILES

`ivtff.py`
: IVTFF 2.0 reader.

`corpus.py`
: Corpus object; selections by locus type, section, language, scribe, page.

`refcorpora.py`
: Reference samples.

`a1_structure.py`
: Codex map, section and quire summaries.

`a2_stats.py`
: Basic statistics, Zipf, word length, entropy.

`a3_wordstructure.py`
: Glyph positions, word rules, slot grammar.

`a4_lines.py`
: Line and paragraph effects.

`a5_groups.py`
: A/B keyness, B-score, group similarity.

`a6_labels_repeats_agreement.py`
: Labels, repetition, transcription agreement.

`a9_decipher.py`
: The experiments of chapter 17.

`build_figures.py, figs_variants.py, vizkit/`
: The figures of Parts I–IV.

`langs.py`
: Builds the 103-language reference samples (and vowel-less forms) from the Christodoulopoulos–Steedman Bible corpus.

`images/paint_map.py`
: Pixel classification of the Yale photographs into pigments.

`build_figures2.py`
: The figures of Part V.

`framework/PROTOCOL.md`
: The pre-registration protocol of chapter 20.

`framework/hypotheses.yaml, results.csv`
: The hypothesis registry (Appendix E) and the log of every test run.

`framework/harness.py`
: Surrogates, held-out split, reference languages, profile statistics.

`framework/t_profile_screen.py, t_subst_screen.py`
: Chapter 21.

`framework/t_sequence.py, t_page_similarity.py, t_form_function.py, t_hmm.py`
: Chapter 22.

`framework/homophonic.py, t_homophonic_control.py, t_homophonic_voynich.py`
: Chapter 23.

`framework/t_register.py`
: Chapter 24.

`framework/t_scribes.py, generators.py, battery.py, t_generators.py`
: Chapter 25: per-scribe tests; the six simulators;

the shared statistics battery; the competition.

`framework/fragment_solver.py, t_fragment_control.py, t_fragment_voynich.py`
Chapter 26.

`framework/t_lists.py; the VOY_TRANSLIT=GC switch in corpus.py and run_gc.sh`
Chapter 27.

`framework/mixedkey.py`
Chapter 28 (with t_mixed_fingerprint.py): the mixed-key simulator, the family map and the registered test.

`framework/t_collapse3.py`
Chapter 29: collapse (with t_collapse.py, t_collapse2.py for the registered and first revised forms and the choice dynamics); t_line_position.py and t_line_position_verse.py, line position; t_simonetta.py, Simonetta's rules.

`framework/elem_solver.py`
Chapter 29: the element solver; useg.py, unit discovery; t_elem_control.py, the ladder; diag_elem.py and diag_seg.py, the two diagnostics; t_elem_voynich.py, the manuscript runs; mixed_solver.py keeps the two objectives that failed.

`build_figures3.py`
The figures of Part VII.

GC2a-n.txt. Glen Claston's v101 transliteration in IVTFF format, version 2a (25 June 2025), distributed at the same address; used in chapter 27 only.

B

Folio Concordance

Every page and fold-out panel in the ZL3b transliteration, in bound order, with its quire, the type of illustration on it (after the transliteration's page variables), Currier's language where he assigned one, Davis's scribe, and the number of words of running text and of labels. Section codes: H herbal, A astronomical, Z zodiac, B bathing (biological), C cosmological, P pharmaceutical, S starred text, T text only.

page	q.	type	lang.	scribe	text	labels
f1r	1	T	A	1	210	0
f1v	1	H	A	1	82	0
f2r	1	H	A	1	80	4
f2v	1	H	A	1	55	0
f3r	1	H	A	1	110	0
f3v	1	H	A	1	85	0
f4r	1	H	A	1	61	0
f4v	1	H	A	1	79	0
f5r	1	H	A	1	54	0
f5v	1	H	A	1	48	0
f6r	1	H	A	1	80	0
f6v	1	H	A	1	113	0
f7r	1	H	A	1	57	0
f7v	1	H	A	1	75	0
f8r	1	H	A	1	134	0
f8v	1	H	A	1	101	0
f9r	2	H	A	1	74	0
f9v	2	H	A	1	74	0
f10r	2	H	A	1	86	0
f10v	2	H	A	1	57	0
f11r	2	H	A	1	59	0
f11v	2	H	A	1	44	1
f13r	2	H	A	1	73	0
f13v	2	H	A	1	61	0
f14r	2	H	A	1	72	0
f14v	2	H	A	1	72	0
f15r	2	H	A	1	78	0
f15v	2	H	A	1	70	0
f16r	2	H	A	1	81	0
f16v	2	H	A	1	76	0
f17r	3	H	A	1	74	2
f17v	3	H	A	1	135	0
f18r	3	H	A	1	79	0
f18v	3	H	A	1	69	0
f19r	3	H	A	1	76	0
f19v	3	H	A	1	78	0
f20r	3	H	A	1	88	0
f20v	3	H	A	1	85	0
f21r	3	H	A	1	104	0
f21v	3	H	A	1	57	0
f22r	3	H	A	1	104	0
f22v	3	H	A	1	71	0
f23r	3	H	A	1	104	0
f23v	3	H	A	1	85	0
f24r	3	H	A	1	109	0
f24v	3	H	A	1	84	0
f25r	4	H	A	1	47	0
f25v	4	H	A	1	57	0
f26r	4	H	B	2	71	0
f26v	4	H	B	2	93	0
f27r	4	H	A	1	85	0
f27v	4	H	A	1	54	0
f28r	4	H	A	1	60	0
f28v	4	H	A	1	63	0
f29r	4	H	A	1	60	0
f29v	4	H	A	1	73	0
f30r	4	H	A	1	92	0
f30v	4	H	A	1	63	0
f31r	4	H	B	2	102	0
f31v	4	H	B	2	110	0
f32r	4	H	A	1	72	0
f32v	4	H	A	1	79	0
f33r	5	H	B	2	74	0
f33v	5	H	B	2	93	0
f34r	5	H	B	2	123	0
f34v	5	H	B	2	117	0
f35r	5	H	A	1	90	0
f35v	5	H	A	1	91	0
f36r	5	H	A	1	53	0
f36v	5	H	A	1	74	0
f37r	5	H	A	1	73	0
f37v	5	H	A	1	89	0
f38r	5	H	A	1	38	0
f38v	5	H	A	1	60	0
f39r	5	H	B	2	157	0
f39v	5	H	B	2	139	0
f40r	5	H	B	2	93	0
f40v	5	H	B	2	99	0
f41r	6	H	B	5	86	0
f41v	6	H	B	5	59	1
f42r	6	H	A	1	136	0
f42v	6	H	A	1	100	0
f43r	6	H	B	2	153	0
f43v	6	H	B	2	151	0
f44r	6	H	A	1	74	0
f44v	6	H	A	1	98	0
f45r	6	H	A	1	89	0
f45v	6	H	A	1	74	0
f46r	6	H	B	2	163	0
f46v	6	H	B	2	110	0
f47r	6	H	A	1	76	0
f47v	6	H	A	1	75	0
f48r	6	H	B	5	89	0
f48v	6	H	B	5	114	0
f49r	7	H	A	1	113	0
f49v	7	H	A	1	131	20
f50r	7	H	B	2	91	0
f50v	7	H	B	2	98	0
f51r	7	H	A	1	86	0
f51v	7	H	A	1	75	0
f52r	7	H	A	1	63	0
f52v	7	H	A	1	77	0
f53r	7	H	A	1	57	0
f53v	7	H	A	1	72	0
f54r	7	H	A	1	99	0
f54v	7	H	A	1	93	0
f55r	7	H	B	2	124	0
f55v	7	H	B	2	94	0
f56r	7	H	A	1	101	0
f56v	7	H	A	1	90	0
f57r	8	H	B	5	83	0
f57v	8	C	–	1	0	5
f58r	8	S	A	3	378	0
f58v	8	S	A	3	374	0

page	q.	type	lang.	scribe	text	labels
f65r	8	H	–	3	0	3
f65v	8	H	–	3	39	0
f66r	8	T	B	5	294	51
f66v	8	H	B	5	113	0
f67r1	9	A	–	4	41	0
f67r2	9	A	–	4	126	64
f67v2	9	C	–	4	33	8
f67v1	9	A	–	4	0	16
f68r1	9	A	–	4	30	29
f68r2	9	A	–	4	36	25
f68r3	9	A	–	4	0	16
f68v3	9	C	–	4	54	2
f68v2	9	A	–	4	43	0
f68v1	9	A	–	4	0	0
f69r	10	C	–	4	36	47
f69v	10	C	–	4	0	0
f70r1	10	C	–	4	0	0
f70r2	10	C	–	4	101	0
f70v2	10	Z	–	4	0	34
f70v1	10	Z	–	4	0	25
f71r	11	Z	–	4	0	17
f71v	11	Z	–	4	0	22
f72r1	11	Z	–	4	0	23
f72r2	11	Z	–	4	0	31
f72r3	11	Z	–	4	0	40
f72v3	11	Z	–	4	0	34
f72v2	11	Z	–	4	0	31
f72v1	11	Z	–	4	0	28
f73r	12	Z	–	4	0	30
f73v	12	Z	–	4	0	30
f75r	13	B	B	2	385	0
f75v	13	B	B	2	312	36
f76r	13	T	B	2	560	9
f76v	13	B	B	2	393	0
f77r	13	B	B	2	321	11
f77v	13	B	B	2	321	10
f78r	13	B	B	2	289	7
f78v	13	B	B	2	294	0
f79r	13	B	B	2	382	0
f79v	13	B	B	2	355	0
f80r	13	B	B	2	434	10
f80v	13	B	B	2	376	0
f81r	13	B	B	2	210	0
f81v	13	B	B	2	252	2
f82r	13	B	B	2	273	13
f82v	13	B	B	2	306	14
f83r	13	B	B	2	340	4
f83v	13	B	B	2	254	4
f84r	13	B	B	2	354	14
f84v	13	B	B	2	341	0
f85r1	14	T	B	2	347	0
f85r2	14	C	B	2	107	0
f86v4	14	C	B	2	47	0
f86v6	14	T	B	2	490	0
f86v5	14	T	B	2	389	0
f86v3	14	C	B	2	277	0
f87r	15	H	A	1	102	0
f87v	15	H	A	1	77	0
f88r	15	P	A	1	133	16
f88v	15	P	A	1	130	13
f89r1	15	P	A	1	124	15
f89r2	15	P	A	1	211	22
f89v2	15	P	A	1	164	15
f89v1	15	P	A	1	145	9
f90r1	15	H	A	1	68	0
f90r2	15	H	A	1	42	0
f90v2	15	H	A	1	61	0
f90v1	15	H	A	1	76	0
f93r	17	H	A	1	148	0
f93v	17	H	A	1	69	0
f94r	17	H	B	3	77	0
f94v	17	H	B	3	87	0
f95r1	17	H	B	3	101	0
f95r2	17	H	B	3	79	0
f95v2	17	H	B	3	55	0
f95v1	17	H	B	3	118	0
f96r	17	H	A	1	78	0
f96v	17	H	A	1	51	0
f99r	19	P	A	1	158	35
f99v	19	P	A	1	137	26
f100r	19	P	A	1	90	18
f100v	19	P	A	1	79	14
f101r	19	P	A	1	227	0
f101v	19	P	A	1	179	18
f102r1	19	P	A	1	97	4
f102r2	19	P	A	1	137	4
f102v2	19	P	A	1	165	20
f102v1	19	P	A	1	117	9
f103r	20	S	B	3	523	0
f103v	20	S	B	3	453	0
f104r	20	S	B	3	448	0
f104v	20	S	B	3	474	0
f105r	20	S	B	3	378	0
f105v	20	S	B	3	399	0
f106r	20	S	B	3	431	0
f106v	20	S	B	3	444	0
f107r	20	S	B	3	487	0
f107v	20	S	B	3	462	0
f108r	20	S	B	3	493	0
f108v	20	S	B	3	580	0
f111r	20	S	B	3	622	0
f111v	20	S	B	3	565	0
f112r	20	S	B	3	398	0
f112v	20	S	B	3	419	0
f113r	20	S	B	3	529	0
f113v	20	S	B	3	501	0
f114r	20	S	B	3	456	0
f114v	20	S	B	3	374	0
f115r	20	S	B	–	461	0
f115v	20	S	B	3	408	0
f116r	20	S	B	3	545	0
f116v	20	T	–	3	0	2

C

Statistical Tables

Values computed from the ZL3b transliteration (running text only, words containing an unreadable character removed) and from reference samples of 34,180 words each. Entropies in bits; the character stream includes the word space as a symbol.

C.1 BASIC STATISTICS

corpus	tokens	types	hapax	mean len.	s.d.	Zipf slope	h_1	h_2
Voynich (all running text)	34,180	7,466	5,266	5.10	1.94	-1.04	3.88	2.13
Voynich A	10,772	3,462	2,530	5.03	2.11	-0.99	3.85	2.19
Voynich B	22,869	4,941	3,427	5.14	1.84	-1.07	3.88	2.01
Latin (Vulgate)	34,180	6,817	4,084	5.30	2.68	-0.95	4.00	3.33
Latin (Pliny, Nat. Hist.)	34,180	12,578	8,976	5.97	2.71	-0.81	4.05	3.42
English (KJV)	34,180	2,351	933	4.08	1.90	-1.17	4.01	3.11
German (Bible)	34,180	3,575	1,634	4.54	2.10	-1.08	4.06	3.15
Italian (Bible)	34,180	4,314	2,286	4.45	2.48	-1.07	4.01	3.20

Voynich figures on the glyph-unit alphabet (ch, sh, cth, ckh, cph, cfh as single symbols): all text h_1 = 3.87, h_2 = 2.26; A h_2 = 2.38; B h_2 = 2.11.

C.2 LABELS

type	description	tokens	types	% q-initial	% o-initial	mean len.	% types in text
Lz	zodiac figure labels	345	281	0.3	67.2	5.66	51
Ln	nymph labels (biological)	64	60	3.1	54.7	5.09	80
Lf	plant-part labels (pharmaceutical)	202	191	0.5	51.0	5.75	52
Lc	container labels (pharmaceutical)	36	36	0.0	52.8	6.94	8
Ls	star labels	86	85	1.2	59.3	5.77	56
Lt	tube/tub labels	55	53	1.8	69.1	5.76	60
L0	other labels	202	133	2.5	22.8	3.52	58
La	astronomical labels	9	9	0.0	55.6	5.78	56
Lp	herb labels	5	5	0.0	20.0	5.20	60
P	running text (reference)	34180	7466	15.5	20.4	5.10	100

C.3 LINE POSITION

Share of words (per cent) by initial or final glyph and by position in the line; 4,034 lines with three or more words.

position	o	c	q	s	d	y	a	p	k	t	f	l
line-first word, initial glyph	13.4	4.1	13.0	14.7	15.4	14.6	0.4	9.3	2.6	9.7	0.8	1.7
line-middle word, initial glyph	21.1	21.3	16.9	12.2	7.1	3.4	5.5	0.5	3.8	1.9	0.3	4.2
line-last word, initial glyph	22.9	16.6	8.8	8.0	16.8	3.5	8.7	0.4	3.0	1.7	0.2	5.9

	y	n	l	r	s	m	o	d	g
line-first word, final glyph	32.4	20.0	14.2	20.4	3.7	0.4	6.0	1.5	0.0
line-middle word, final glyph	42.2	16.2	16.5	15.2	3.0	0.9	3.2	1.4	0.1
line-last word, final glyph	39.0	15.3	11.1	8.6	3.4	15.6	1.3	2.6	2.2

C.4 WORDS MOST CHARACTERISTIC OF A AND OF B

Counts and rate per 10,000 words; ranked by signed log-likelihood G^2.

word	A	B	/10k A	/10k B
daiin	456	296	423	129
chol	261	99	242	43
chor	165	27	153	12
cthy	88	5	82	2
sho	94	14	87	6
cthol	51	2	47	1
chy	97	33	90	14
s	161	98	149	43
shol	107	56	99	24
cthor	38	2	35	1
cho	55	16	51	7
shor	64	26	59	11
qotchy	48	13	45	6
ckhey	27	1	25	0
otchol	25	1	23	0
kchy	29	4	27	2
shy	56	28	52	12

word	A	B	/10k A	/10k B
chedy	1	486	1	213
shedy	2	419	2	183
qokeedy	0	303	0	132
qokedy	1	267	1	117
qokain	4	274	4	120
otedy	0	138	0	60
qokaiin	22	243	20	106
okedy	0	106	0	46
lchedy	1	115	1	50
okeedy	0	103	0	45
qotedy	0	87	0	38
qol	6	132	6	58
qoteedy	0	76	0	33
oteedy	1	86	1	38
qokal	16	168	15	73
qokeey	40	262	37	115
okain	8	125	7	55
sheedy	1	81	1	35
chdy	8	122	7	53
otain	3	90	3	39
keedy	0	62	0	27
al	24	183	22	80
otar	8	113	7	49
aiin	81	366	75	160
ol	99	419	92	183

C.5 SUBSTITUTION SEARCH

Mean log-probability per character (natural log) under each language's bigram model. In every case the control (a random encipherment of the language's own text, solved by the same search) was recovered with a score equal to the own-text score. "Hits" is the share of mapped Voynich words found in the reference sample's vocabulary, under the best key and under a random key.

language	own text	best key	random key	hits, best	hits, random
Latin (Vulgate)	-2.307	-2.477	-4.744	7.9%	0.6%
Latin (Pliny, Nat. Hist.)	-2.370	-2.437	-4.552	8.2%	2.5%
English (KJV)	-2.154	-2.443	-4.505	6.9%	1.1%
German (Bible)	-2.184	-2.516	-5.735	3.8%	0.3%
Italian (Bible)	-2.221	-2.501	-5.363	8.8%	1.8%
Latin (Vulgate), vowel-less	-2.131	-2.426	-3.487	28.4%	8.8%
Latin (Pliny, Nat. Hist.), vowel-less	-2.222	-2.398	-3.603	38.1%	11.1%
English (KJV), vowel-less	-1.968	-2.559	-4.366	19.2%	5.9%
German (Bible), vowel-less	-2.046	-2.563	-4.138	20.5%	8.7%
Italian (Bible), vowel-less	-2.096	-2.477	-4.430	19.8%	6.7%

D

How This Book Was Made

This book was written by a human author working with an AI system, Claude, made by Anthropic. The reader is entitled to know what that means in practice, and what it does not.

WHAT EACH DID

The author set the question, the standard of evidence and the shape of the book, chose the format and the figure style from alternatives, and is responsible for every claim in it. The AI system did the computation: it fetched the transliterations and reference texts, wrote and ran the programs in Appendix A, built the figure library, drafted the chapters from the results and the published literature, and checked the draft against the data. Where the draft said something the data did not support, it was changed or cut. Where the literature was uncertain, the uncertainty is in the text.

WHAT WENT WRONG FIRST

An earlier draft of this project, made in the spring of 2026, was built from the manuscript images alone, without a machine-readable transcription. The AI system of that draft produced "readings" of labels and words by describing what it saw in the pictures, and from those readings a theory grew: names of zodiac signs, names of planets, a seven-planet pharmacology, a constructed language. When the readings were checked against the transliterations in October 2026, none of them existed. The strings the system had reported were not in the manuscript in any transcription; several contained sequences the script cannot form. The earlier draft was set aside and this book was begun from the data.

We record this because it is the characteristic failure of the subject, and because an AI system is unusually good at committing it. A language model asked to read an unreadable page will read it, fluently and with confidence, and the result will have every outward feature of a finding. The only protection is the one this book has tried to apply throughout: no number without a program that produces it, no reading without a transcription behind it, no claim about the object without a source that examined the object. Readers who find places where that standard slips are asked to say so.

REPRODUCIBILITY

The transliterations, the reference texts, the programs and the figure sources are distributed with the book. Every figure can be regenerated from them, and every number in the text can be traced to a table those programs write. The folio images are the Beinecke Library's own and are free to use.

E

Hypothesis Registry

The registry as it stood when the book went to press. Each entry records the hypothesis, the predictions written before testing, the tests run, the result, and the status under the decision rules of chapter 20. Full test outputs are in `data/framework/out/`. The entry for H-CIPHER-MIXED (Part VII) was frozen before any of its tests were run, in `out/hypotheses_preregistered_partVII.yaml`, SHA-256 `803e20975abdb5642d322d399eea0aa5-0638e20353d1fa5e1eb8b7f8f9cd1c66`.

H-LANG-ANY LANG

Statement. The text is a natural language written plainly in an alphabet (one symbol ≈ one sound).

Predictions. (1) character-level statistics (h1, h2, word length, short-word share) fall inside the range of attested languages (2) word-level statistics (TTR, hapax, Zipf, top-word share) fall inside the attested range

Tests. T1-profile

Result. h2 = 2.13 (2.26 glyph units) below the minimum of 103 languages in plain and vowel-less form (2.51); all other 13 statistics inside the 95 % range; nearest languages differ by level (character level: Malagasy, Akawaio, Maori, Cebuano — CV-syllable languages; word level: Quechua, Swahili, Zulu, Nahuatl — agglutinative).

Status. not supported (plain alphabet); inconclusive for syllabic or positional scripts

H-LANG-SUBST LANG/CIPHER

Statement. The text is one of 103 languages under a one-to-one glyph-to-letter substitution (with or without vowels).

Predictions. (1) the best key, fitted on training quires, scores on held-out quires within the planted-key control's tolerance of the language's own text (2) dictionary coverage of decoded held-out words exceeds random keys by more than the control margin

Tests. T2-subst-screen

Result. see out/subst_screen.csv (187 runs); no language meets both criteria

Status. not supported

H-CIPHER-HOMOPHONE CIPHER

Statement. The text is a known language under a homophonic/verbose cipher (several ciphertext units per plaintext letter), units being glyphs, slot-fillers, BPE tokens or whole words.

Predictions. (1) one-glyph variants of common words should behave alike in context (homophone substitution does not change the plaintext) — form/function coupling above any natural language (2) exact phrase repetition should be rare even where word repetition is high (homophone choice breaks phrase identity) (3) a homophonic solver validated on planted keys should, for the right language and unit, reach held-out 4-gram scores near the language's own and dictionary coverage well above random

Tests. T6-form-function, T4d-phrases, T3a-control, T3b-homophonic
Result. T6 — Spearman(form, context) 0.147 and 15 % of nearest context-neighbours within edit-1, vs $\leq$0.09 / $\leq$9 % for 21 language samples incl. vowel-less and abjads; T4d — zero repeated 4-word phrases (Pliny 16, Bible texts 150–900); T3a — solver recovers planted 80-unit Latin keys (100 % with spaces, 98 % without; vowel-less Latin 98 %, gap 0.02); T3b — 19 searches (Latin, Italian, German, French, Czech, vowel-less Latin $\times$ glyph / slot / merged-pair / word units $\times$ spaces kept or removed): held-out gap 0.7–3.5 nats for all plain languages (control 0.02), keys collapsed onto 8–17 letters, dictionary rates largely reproduced by frequency-matched permutations of the found keys; vowel-less Latin gap 0.34, 82 % dictionary vs 57 % permuted, output not Latin. Re-run with the stronger frequency penalty (lam = 2) and polishing sweep of ch. 26: plain-language gaps 2.0-2.5 unchanged; vowel-less Latin 0.72-0.73 (the earlier 0.34 was the alphabet collapsing onto a few consonants)
Status. two structural predictions supported; the solver finds no key for any plain language or unit; the vowel-less near miss was an artefact of alphabet collapse

H-CIPHER-WORDLETTER CIPHER
Statement. Each Voynich word encodes one plaintext letter (Naibbe-type), word spaces are not plaintext spaces.
Predictions. (1) word sequence treated as a letter sequence has conditional entropy near a language's letter entropy (3.0–3.5 bits) once homophones are merged (2) words300 homophonic search (no spaces) should find a key with dictionary coverage above random
Tests. T3b-homophonic words300
Result. Latin, Italian and German with the 300 commonest words as units (spaces removed): held-out gap 2.5–3.5 nats; the solver assigns nearly all units to two or three letters; no words of four letters or more in the dictionary
Status. not supported in this form (one word = one letter, fixed table); fragmenting ciphers with learned segmentation remain untested

H-MEANINGLESS-AUTOCOPY MEANINGLESS
Statement. The text was generated by copying and modifying nearby words (Timm & Schinner).
Predictions. (1) near-twin rate should depend strongly on distance and collapse when words are shuffled within a page (2) words that look alike should behave alike (copying preserves context) — consistent with T6 (3) no long-range (section-level) vocabulary structure beyond what local copying produces
Tests. E4b (book ch. 17), T6, T5
Result. shuffling within page removes only 1–3 points of the near-twin rate (24 vs 27 % within ten words); T6 supports; section/sheet structure (T5, contrasts) is strong and not explained by local copying without a steering rule
Status. inconclusive — predictions 2 supported, 1 and 3 not

H-STRUCT-PHRASES STRUCT
Statement. The text contains repeated formulaic phrases, as recipe and herbal texts do.
Predictions. (1) repeated 3- and 4-word sequences at rates comparable to Latin prose
Tests. T4d
Result. 97 repeated 3-grams, 0 repeated 4-grams in 34k words; Pliny 128/16; Bible texts 700–3000 / 170–1700
Status. not supported (anomaly recorded)

H-STRUCT-LINE STRUCT
Statement. The line is a unit of the system (not an accident of layout).
Predictions. (1) line-initial/final glyph distributions differ from line-internal; MI beyond the word is confined to the line
Tests. a4_lines (book ch. 7), T4a
Result. supported (book ch. 7); MI at d=10 is 0.015 in the text, 0.012 with words shuffled within lines, 0.009 with words shuffled across pages
Status. supported

H-ORDER-SHEET ORDER
Statement. The sheet (bifolio) is the unit of composition; pages on one sheet share content.
Predictions. (1) same-leaf and same-sheet page pairs are more similar in vocabulary than random pairs of the same section and language
Tests. T5-page-similarity
Result. z = 9.6 (same leaf), 9.3 (same sheet), permutation null
Status. supported

H-ORDER-BINDING ORDER
Statement. The present binding order preserves the original sequence of sheets.
Predictions. (1) pages bound next to each other on different sheets are more similar than random pairs
Tests. T5-page-similarity
Result. z = 1.0, p = 0.15 (adjacent, different sheet); z = 0.75 at 2–3 pages
Status. not supported — independent corroboration of Layfield & Davis (2026)

H-STRUCT-WORDCLASS STRUCT
Statement. Common words fall into distributional classes (function-word-like vs content-like) as in languages.
Predictions. (1) context clustering of the 300 commonest words yields classes not reducible to spelling families
Tests. T6 (nn5_edit1), planned: Brown clustering
Result. nearest context neighbours are dominated by spelling variants (15 %), i.e. classes largely coincide with form
Status. inconclusive

H-IMAGE-PAINT IMAGE
Statement. Paint was applied in campaigns organised by section/quire rather than by scribe.
Predictions. (1) pigment coverage varies by section and quire more than by scribe within section

Tests. I1-paint-map
Result. green confined to herbal/bathing; blue on 22 pages scattered across quires 1–7, 9, 13; no scribe effect within section detectable at 2,000 px
Status. inconclusive (data layer built)

H-IMAGE-ZODIAC-LABELS IMAGE
Statement. Zodiac labels are unique names of degrees or stars.
Predictions. (1) labels unique across signs; no shared sign-level marker
Tests. a6_labels (book ch. 11), zodiac figure
Result. 37 of 281 labels recur across signs; ot-/ok- drift from early to late signs
Status. not supported

H-STRUCT-RING STRUCT
Statement. Text written in rings round diagrams, and radial text, is written in the same register as paragraph text (same word-initial and word-final distributions).
Predictions. (1) ring and radial words begin with q at the paragraph rate (about 15 %) and with o at the paragraph rate (about 20 %) (2) on the same pages
Tests. T8-register
Result. q-initial: paragraph 15.5 %, ring 1.9 %, radial 4.9 %, label 1.0 %; o-initial 20 / 40 / 35 / 54 %. On the seven pages carrying both ring and paragraph text, the paragraph text has 6.5 points more q-initial words (permutation $p < 0.001$). Ring text is in the label register, not the paragraph register.
Status. not supported (ring text is label-like)

H-CIPHER-NULLS CIPHER
Statement. One or two glyph units are nulls (meaningless insertions); the remaining text is a plain alphabet.
Predictions. (1) deleting the null unit(s) raises conditional character entropy into the attested-language range (at least 2.51 bits)
Tests. T9-nulls
Result. base h2 2.26 (glyph units); best single deletion raises it to 2.40 (o), best pair to 2.40 (o+a); 103-language minimum 2.51. Twelve languages change by at most 0.1 bits when their commonest letter is deleted.
Status. not supported

H-CIPHER-ANAGRAM CIPHER
Statement. Each word is a set of letters written out in one fixed canonical order (an anagram cipher with sorting), so that within-word order carries no information.
Predictions. (1) nearly all words are monotone under one global ranking of the units; languages are far lower
Tests. T10-order
Result. 44 % of words are monotone under the best single ranking (glyph units; 41 % EVA letters), vs 9 % with letters shuffled within words and 10-32 % for twelve languages (Greek 32 %, vowel-less Latin 31 %, Latin 19 %, Hungarian 10 %).
Status. not supported in the strict form; order is more predictable from the letter set than in any language sampled, as the slot grammar implies

H-LANG-WORDLEVEL LANG/CIPHER
Statement. Whatever the symbols are, the word sequence carries information the way a language's does: knowing the previous words reduces the cost of the next by as much as it does in attested languages (the cross-linguistic law of Bentz et al. 2017 relating unigram entropy to entropy rate).
Predictions. (1) the manuscript's entropy rate lies within 2 sd of the line fitted on 103 languages at its unigram entropy; the gain from word order (compressed size of the shuffled text minus the text) lies inside the language range
Tests. T12-entropy-law
Result. Compression estimator (random word ids): law refit on 102 equal-length samples with word spaces, $r = 0.98$; the manuscript lies 3.7 sd above the line (its words are harder to predict than a language with its vocabulary would be), its word-shuffled form 4.5 sd. Gain from word order 0.42 bits per word, against 1.34 (the lowest language) to 4.10, median 2.39; no language is within three times of it. Japanese excluded: its sample has no word spaces, so its 'words' are phrases.
Status. not supported — the word sequence carries less information than any attested language's; consistent with a cipher that destroys or disguises word order, with a list-like text, or with a generated text

H-CIPHER-TRANSPOSITION CIPHER
Statement. The glyph stream has been transposed with a fixed period (columnar, route or decimation), as the Zodiac-340 cipher was.
Predictions. (1) the count of repeated bigrams when the stream is read at period p peaks sharply at the true period (2) and the periodic coincidence kappa(p) shows a peak absent from the within-word shuffle
Tests. T11-periodicity
Result. Period-bigram counts are flat: the best period ($p = 2$) exceeds the median over $p = 2..60$ by 0.04 %, in the text and in every surrogate alike. kappa(p) peaks at $p = 6$, the mean word length plus space, in the text and in all surrogates that keep the words.
Status. not supported

H-CIPHER-POLYALPHABETIC CIPHER
Statement. The substitution changes with position in the line (a line-keyed or position-keyed alphabet), so that the same plaintext at the same offset in different lines appears as glyph-disjoint isomorphs.
Predictions. (1) consecutive lines coincide at matched positions more than random line pairs (2) beyond what shuffling words within lines preserves; glyph-disjoint isomorph pairs occur at matched line positions more often than in the within-line word shuffle
Tests. T11-line-kappa, T11-isomorphs
Result. Consecutive-line coincidence 10.5 % vs 9.4 % for random pairs; the words-in-lines shuffle keeps 9.8 vs 9.1 %, so the small excess is neighbouring lines sharing words, not positions. Glyph-disjoint isomorphs at matched

positions: 2.5 % of all such pairs in the text, 2.4 % in the within-line shuffle.
Status. not supported

H-STRUCT-CONTEXT STRUCT
Statement. Predictive information reaches beyond the word: a model that sees more of the preceding text predicts the next unit better, on sheets it was not trained on, as it does in languages.
Predictions. (1) held-out bits per glyph keep falling as the model's context grows past one word (order 6-8); held-out bits per word fall when the previous word is known (order 2 beats order 1) (2) by an amount inside the language range
Tests. T13-context-scaling
Result. Glyph units, Kneser-Ney, trained on the training quires and scored on the held-out quires: 3.89 bits at order 1, 2.44 at order 2, 2.24 at order 4 (the minimum), 2.50 at order 8; the gain from order 2 to the best order is 0.20 bits, against 1.42 (Latin), 1.19 (Italian), 1.27 (German), 0.97 (Hebrew). Words: 10.87 bits at order 1 and 11.02 at order 2 - the previous word makes the held-out prediction worse - against gains of 0.3 to 1.2 bits in eight languages. The lines-shuffled surrogate is slightly more predictable than the text at every order because the split is then no longer by quire.
Status. not supported: nearly all predictive information is inside the word; the previous word carries none that transfers to other sheets

H-HANDS-PROCEDURE PROCESS
Statement. The scribes composed the text by executing a procedure individually, leaving hand-specific traces: each scribe's own bias in choosing between variant spellings, and his own transition and error profile.
Predictions. (1) within one section and language (2) scribes differ in the share of each variant (ch/sh (3) k/t (4) o/a (5) e/ee (6) y/dy (7) in/iin (8) ain/aiin (9) l/r (10) d/s (11) ol/or) beyond a null that shuffles scribe labels across pages; glyph-bigram and near-twin-substitution profiles differ likewise
Tests. T14-scribes
Result. Herbal B, scribes 2, 3 and 5 (3,316 words on 32 pages): variant choice G = 54.6, z = 1.2, p = 0.09 against the page-shuffled null - the three hands pick the same variants at the same rates (ch/sh 0.66/0.64/0.59; y/dy 0.37/0.41/0.37; d/s 0.67/0.71/0.75); glyph-bigram divergence z = 6.9; near-twin substitution profile z = 2.0, p = 0.02; visual-confusion share of near-twin substitutions 0.17/0.17/0.18. The section control (scribe 1, herbal vs pharmaceutical) gives z = 19 / 17 / 3.3 on the same statistics: section outweighs hand by an order of magnitude.
Status. not supported for variant choice (hands interchangeable); weakly supported for transition profile. Consistent with a fixed table executed identically, or with copying; not with free individual composition. Power limited by 500-word samples for scribes 3 and 5.

H-TRANSLIT-ROBUST STRUCT
Statement. The results of Part V do not depend on the transliteration system: they hold for Glen Claston's v101 (GC2a, one character per glyph shape, 69 symbols) as for ZL (EVA, 27 letters / 29 glyph units).
Predictions. (1) each registered statistic computed on GC falls on the same side of its decision rule as on ZL
Tests. T18-translit (the battery re-run with VOY_TRANSLIT=GC)
Result. Word-level results hold unchanged: TTR 0.229 (ZL 0.218), hapax 0.70 / 0.71, Zipf -1.01 / -1.04, word-order gain residual 2.8 sd (ZL 2.9 before, 3.7 with random ids), held-out word bigram worse than unigram (11.02 → 11.17 bits), form-function 0.16 / 24 % (ZL 0.15 / 15 %), ring register q-initial 1.8 % vs 14.5 % in paragraphs, 3 repeated 4-grams (ZL 0), nearest-language families the same, periodicity and isomorph tests flat, sorted-anagram share 54 % (ZL 44 %). Glyph-level results move: conditional character entropy is 2.53 bits under GC against 2.13 under ZL, i.e. at the attested minimum (2.50) rather than far below it, and deleting the two commonest GC symbols lifts it to 2.75, inside the language range, so the null-glyph exclusion of T9 does not hold under GC.
Status. supported for every word-level result; the symbol-level entropy anomaly and the null-glyph exclusion are transliteration-dependent and are reported as such

H-TEXT-LIST LANG
Statement. The plaintext is a list-like text (recipes, an index, a glossary, a star catalogue), whose items carry little word order, so that the manuscript's weak word order is a property of the genre rather than of the encoding.
Predictions. (1) Latin texts of those genres show word-order gains and held-out bigram gains near the manuscript's at matched length (2) and repeat phrases as rarely
Tests. T17-lists
Result. At matched lengths the manuscript gains 0.27-0.37 bits per word from order and loses 0.13-0.16 bits on held-out pages when the previous word is used. Apicius (recipes) gains 2.32 / +0.91; Pliny's index 2.54 / +1.37; Cato 1.40 / +0.51; Hyginus 1.22 / +0.64; Isidore's glossaries 0.63-0.78 / +0.08-0.31; Pliny's prose 0.59 / +0.05; the Vulgate 1.45 / +0.65. Repeated 4-word phrases: 61-313 in the recipe, index and star texts, 6-18 in the glossaries and prose, 0-1 in the manuscript. The list genres are more formulaic than prose, not less; the nearest Latin genre by order gain is the encyclopedic glossary, and it still differs in sign on the held-out test.
Status. not supported: no Latin genre sampled has the manuscript's combination of near-zero order and near-zero phrase repetition

H-PROCESS-GENERATORS PROCESS
Statement. One of the mechanical procedures proposed for the text - independent slot-grammar sampling, Timm-Schinner self-copying, Rugg's table and grille, a Naibbe-type affix-table cipher of Latin (plain or vowelless), a Tranchedino-type homophonic cipher with nulls and re-spacing, a Trithemius-type word-for-letter table - reproduces the manuscript's held-out statistics once its free parameters are fitted on the training quires.
Predictions. (1) for the true procedure (2) the manuscript's value of each held-out statistic (word-order gain (3) held-out bigram gain (4) glyph context gain (5) form-function coupling (6) repeated 3- and 4-grams (7) near-twin rates (8) page locality (9) edge-glyph mutual information (10) line-initial divergence) falls inside the spread of 20 simulated corpora written into the

manuscript's own layout

Tests. T15-generators

Result. No generator passes more than two of twelve held-out statistics. The affix-table cipher with large tables (60 prefixes and 60 suffixes per unit, 60 % bigram units) is nearest on the fitted statistics (h2 2.63, TTR 0.24, hapax 0.58) and reproduces the weak word order (gain 0.00 vs 0.40; held-out bigram -0.31 vs -0.12), but has no page locality (z 1.9 vs 16), no form-function coupling (0.04 vs 0.15) and no edge-glyph dependency (0.008 vs 0.21 bits). Self-copying reproduces near-twin rates and the nearest-neighbour share but over-generates order (2.9 bits) and lacks locality. The re-spaced homophonic stream alone reproduces the edge-glyph dependency (0.22). No procedure reproduces the line-initial effects, which all lack by construction. Languages laid out in the same pages give page locality z 10-20 and edge MI 0.04-0.29, both of which the manuscript matches.

Status. not supported for any single procedure as specified; the manuscript combines a verbose cipher's word statistics with a language's topical locality and a coupling of spelling to use that nothing simulated produces

H-CIPHER-FRAGMENT CIPHER

Statement. The written words are fragments of a verbose cipher of a known language: each word is one or more segments of glyphs, each segment encoding one plaintext unit, with the segmentation unknown and plaintext spaces not written.

Predictions. (1) a solver that learns the segmentation jointly with the key (2) validated on planted affix-table ciphers (3) decodes held-out sheets to text that scores within the control's margin of the language and is covered by dictionary words at the control's rate

Tests. T16a-control, T16b-fragment

Result. Control: planted Latin affix-table keys (four prefixes and four suffixes per letter) with one plaintext letter per written word are recovered at 81 % held-out word accuracy (chance 13 %; vowel-less Latin 87 %); with two letters per written word the solver fails (5 %, chance 2 %) - its limit. Manuscript, paragraphs: Latin gap 0.78 nats against 1.53 for permuted keys allowed the same re-segmentation, Italian 0.91 vs 1.49, vowel-less Latin 0.35 vs 0.38; dictionary coverage of the decoded text 33, 33 and 72 % against 32, 31 and 66 % for permuted keys (the control decodes at 51 % and 79 %); rings: Latin 1.04 vs 1.78, coverage 32 vs 26 %. The decoded text is Latin-like letter sequence without Latin words (aeietisciielimeaeeiantiasutanitatiinus...). Segmentations found: 1.5 segments per written word in paragraphs, 2.0 in rings, with y, ol, o, ch, l, dy, qo, or, r, daiin, ar, aiin, edy, chedy, shedy as the commonest segments.

Status. not supported for one plaintext unit per written word (the form the solver can recover); untested for two or more units per word, where the solver fails its own control

H-CIPHER-MIXED CIPHER

Statement. The text is Latin (or an Italian vernacular) enciphered with an Italian chancery mixed key of the Tranchedino type: several glyph-string signs per letter (more for vowels), signs for common syllables, a nomenclator of whole-word signs, and nulls; written either one plaintext word per written word, or with plaintext words split across written words, or as a continuous stream re-spaced.

Predictions. (1) P1 fingerprint: a simulator of this family, fitted on the training quires, reproduces both the weak word order (order gain within 0.25 of 0.42; held-out bigram gain <= 0) and the topical page locality (z within 2 of 16) that no earlier procedure reproduced together, and the form-function coupling (rho >= 0.10) (2) P2 collapse: merging spelling variants that behave alike in context (edit-1 pairs with high context similarity) raises the manuscript's word-order gain and repeated 4-grams by more than merging the same number of random edit-1 pairs, and by more than the same procedure does to Latin; on the simulator the context-guided collapse recovers part of the oracle (true-plaintext) order gain (3) P3 choice dynamics: successive occurrences of the same variant family within a page show a same-variant rate different from the page-shuffled expectation (below it = a human cycling homophones; above it = sticky choice or copying; equal = random choice), and the rate is the same for each scribe (4) P4 solver: a mixed-unit solver with nulls that recovers planted keys of this family on Latin at >= 50 % held-out word accuracy (or >= 70 % of the plaintext's dictionary coverage) decodes the manuscript's held-out sheets to dictionary coverage above its own permutation null by more than the control's margin; if the solver fails its controls at some rung of the ladder, the hypothesis is reported as untested beyond that rung

Tests. T19-mixed-fingerprint, T20-collapse, T21-choice-dynamics, T21c-line-position, T22-mixed-solver, T23-simonetta-rules

Result. P1: 216 keys mapped over the documented range (homophones 2-6/1-4, 60-300 syllables, nomenclator 0-300, nulls 0-10 %) and three spacing practices; none within 2 of the manuscript's page locality (max 10.2, bounded by the plaintext's 10.3); fitted keys (20 corpora each) fail 12 (word spacing), 10 (split words) and 9 (stream) of 12 held-out statistics; line-initial effect 0.001-0.002 against 0.198. P2 (revised to sign-level alternations at matched vocabulary after the registered word merge chained): context merges give about three times the repeated 4-grams of random merges at equal vocabulary but no return of word order (0.27 bits after vs 0.40 before); Latin shows part of the same effect; the simulated keys give too few single-glyph alternations for the test to see them. P3: same-variant ratio 1.04 (scribes 1.02-1.08), but the measure does not separate random, cycling and sticky choice on the simulated keys. Added (T21c): within minimal-pair frames, the variant written depends on line position for 6 of 10 alternations (z 3.9-9.2), none for Latin laid out in the same lines or one verse per line (enciphered or not) or the simulated keys (1 of 33 at z 3.4). P4: element-vocabulary solver (after Lasry 2024) recovers planted keys with known sign boundaries (key 98/97/95/86 %, held-out words 89/86/81/53 % for letters, + nulls, + syllables, + nomenclator) and fails when the signs must be found among the glyphs (unit discovery: 25 % of key, 9 % of words; no word spaces 20 %); the joint loop moves away from the truth even when started near it. Manuscript (exploratory): real spelling beats permuted spelling by 0.18-0.27 nats/element (Latin), 0.44 (Italian), against 0.50-0.60 for the planted cipher and 0.00-0.02 for content-free generators; no Latin words

(coverage 27 % vs 29 % permuted).
Status. not supported in its simulated forms (P1, P2); untested beyond rung L4 of the solver ladder, i.e. as a cipher whose sign boundaries are unknown (P4, by the registered rule); P3 not measurable with the instrument. The line dependence of variant choice means the text as written is not the direct output of a key used in the ordinary way.

F

Methods Inventory

Six surveys were made for the second and third phases of this work: cryptanalysis; computational linguistics and datasets; bioinformatics and signal analysis; imagery and symbology; recent mathematical and machine-learning methods; and, for Part VI, the table-driven encoding procedures of 1380–1520 and the statistics that fingerprint a mechanical procedure. Each inventoried the techniques of its field that could be applied to a 37,000-word symbol sequence or to 213 photographs, recorded whether and how each had been applied to the manuscript, and ranked the untried ones. The surveys, with their verified and unverified citations marked, are distributed with the book in `analysis/research_reports/`. This appendix condenses them into one table of what has been done, by whom, and what has not. "Here" means in this book, with the chapter; "elsewhere" means in the published literature, with the main source; "untried" means we found no application to the manuscript.

Cryptanalysis

technique	status
one-to-one substitution, many languages	here (ch. 17, 21); Hauer & Kondrak 2016
homophonic / verbose solver with planted-key control	here (ch. 23, re-run with the stronger frequency penalty in ch. 26); no prior application found
joint segmentation and substitution (fragmenting cipher)	here (ch. 26), with planted affix-table controls; untried before
frequency-matched permutation null for found keys	here (ch. 23)
word-as-letter	here (ch. 23); Greshko 2025 built the

Linguistics

technique	status
character and word entropy, many languages	here (ch. 9, 21); Bowern & Lindemann 2021
profile against 103 languages, by level	here (ch. 21)
entropy law, unigram vs rate (Bentz et al. 2017)	here (ch. 23); untried before
held-out context scaling (n-gram orders)	here (ch. 23)
vowel/consonant induction (Sukhotin; HMM)	here (ch. 22); Reddy & Knight 2011 on two pages
slot grammar	here (ch. 6); Stolfi, Zattera 2022
form–function coupling of word variants	here (ch. 22); untried before
morphology induction (Morfessor, adaptor grammars)	untried; Bowern & Lindemann note it
word-class induction with held-out gain (Brown, information bottleneck)	untried in the held-out form
Menzerath–Altmann, law of abbreviation	untried
transliteration dependence (ZL vs GC/v101)	here (ch. 27): word-level results hold, symbol-level entropy moves
list-genre reference texts (recipes, index, glossary, star list)	here (ch. 27)

Sequence analysis from other fields

technique	status
mutual information at a distance	here (ch. 22); Montemurro & Zanette 2013
compressibility, BPE merge curve	here (ch. 22)
repeated passages	here (ch. 22)
page similarity vs physical structure	here (ch. 22); Layfield & Davis 2026 by another measure
local alignment of line pairs (Smith–Waterman)	untried
Bayesian word segmentation of de-spaced text	untried
Zipf-preserving long-range surrogates	untried; Montemurro & Degli Esposti 2026
grammar induction (Sequitur, MDL)	untried
change-point detection on line features	untried
burstiness and recurrence intervals	untried
simulation-based competition of generators (six procedures, fitted, scored on held-out statistics)	here (ch. 25)
per-scribe variant-choice and error-profile tests	here (ch. 25)
edge-glyph mutual information; line-initial divergence	here (ch. 25, in the battery)

Images

technique	status
pigment map from photographs	here (ch. 24)
register of ring, radial and label text	here (ch. 24)
zodiac table	here (ch. 24); Zandbergen, Ponzi 2016
scored comparison of zodiac cycles	Ponzi 2016, 131 cycles; not yet as data
paint-transfer (offset) matrix against page order	untried
figure-attribute database (c. 550 figures)	untried
star census of the recipes section	untried
plant duplicates, herbal against pharmaceutical fragments	untried; one pair known
water-channel continuity across bathing pages	untried; Pelling's f78v–f81r claim
glyph-shape clustering from images against EVA	untried
multispectral imaging	ten folios, 2014; published 2024

Bibliography

Nada Aldarrab and Jonathan May. Segmenting numerical substitution ciphers. arXiv:2205.12527, 2022.

Diego R. Amancio, Eduardo G. Altmann, Diego Rybski, Osvaldo N. Oliveira, and Luciano da F. Costa. Probing the statistical properties of unknown texts: application to the Voynich manuscript. *PLoS ONE*, 8(7):e67310, 2013.

Beinecke Rare Book and Manuscript Library. Cipher manuscript (Voynich manuscript). Beinecke MS 408. Yale University Library Digital Collections, `https://collections.library.yale.edu/catalog/2002046`; catalogue description at `https://pre1600ms.beinecke.library.yale.edu/docs/pre1600.ms408.HTM`, 2020. Accessed October 2026.

William Ralph Bennett. *Scientific and Engineering Problem-Solving with the Computer*. Prentice-Hall, Englewood Cliffs, 1976. Chapter 4 computes character entropies for the Voynich text.

Christian Bentz, Dimitrios Alikaniotis, Michael Cysouw, and Ramon Ferrer-i Cancho. The entropy of words—learnability and expressivity across more than 1000 languages. *Entropy*, 19(6):275, 2017. doi: 10.3390/e19060275.

Claire L. Bowern and Luke Lindemann. The linguistics of the Voynich manuscript. *Annual Review of Linguistics*, 7:285–308, 2021. doi: 10.1146/annurev-linguistics-011619-030613.

Christos Christodoulopoulos and Mark Steedman. A massively parallel corpus: the Bible in 100 languages. *Language Resources and Evaluation*, 49(2):375–395, 2015. doi: 10.1007/s10579-014-9287-y.

Raymond Clemens, editor. *The Voynich Manuscript*. Yale University Press, New Haven, 2016. With an introduction by Deborah Harkness and essays by René Zandbergen, Arnold Hunt, Jennifer Rampling, Alain Touwaide, Greg Hodgins, and Paula Zyats and colleagues.

Giorgio Costamagna. Un'ottima applicazione quattrocentesca del sistema cifrante monoalfabetico. *Studi di Storia Medioevale e Diplomatica*, 2:353–358, 1977.

Prescott H. Currier. Some important new statistical findings. In M. E. D'Imperio, editor, *New Research on the Voynich Manuscript: Proceedings of a Seminar, 30 November 1976*. Washington, D.C., 1976. Reprinted in D'Imperio (1978) and at voynich.nu.

Lisa Fagin Davis. Voynich manuscript "solution" debunked. Statement on behalf of the Medieval Academy of America, responding to Nicholas Gibbs, *Times Literary Supplement*, 5 September 2017, 2017.

Lisa Fagin Davis. How many glyphs and how many scribes? Digital paleography and the Voynich manuscript. *Manuscript Studies: A Journal of the*

Schoenberg Institute for Manuscript Studies, 5(1):164–180, 2020.

Lisa Fagin Davis. Multispectral imaging of the Voynich Manuscript: The Lazarus Project images. Manuscript Road Trip (blog), 2024. Ten folios imaged in 2014; processed images released 2024.

Lisa Fagin Davis. Voynich codicology. Manuscript Road Trip (blog), https://manuscriptroadtrip.wordpress.com/2025/01/19/voynich-codicology/, January 2025.

M. E. D'Imperio. *The Voynich Manuscript: An Elegant Enigma*. National Security Agency, Center for Cryptologic History, Fort George G. Meade, 1978.

John Goldsmith and Aris Xanthos. Learning phonological categories. *Language*, 85(1):4–38, 2009. Article in *Language* 85(1).

Michael A. Greshko. The Naibbe cipher: a substitution cipher that encrypts Latin and Italian as Voynich manuscript-like ciphertext. *Cryptologia*, 2025. doi: 10.1080/01611194.2025.2566408. Published online 26 November 2025.

Bradley Hauer and Grzegorz Kondrak. Decoding anagrammed texts written in an unknown language and script. *Transactions of the Association for Computational Linguistics*, 4:75–86, 2016.

Jürgen Hermes. Polygraphia III: The cipher that pretends to be an artificial language. In *Proceedings of the International Conference on the Voynich Manuscript 2022 (CEUR Workshop Proceedings, vol. 3313)*, 2022. URL https://ceur-ws.org/Vol-3313/paper7.pdf. Trithemius' word-for-letter tables measured against the manuscript.

Greg Hodgins. Forensic investigations of the Voynich manuscript. In Raymond Clemens, editor, *The Voynich Manuscript*. Yale University Press, New Haven, 2016. Radiocarbon dating carried out at the University of Arizona, 2009.

Arnold Hunt. Voynich the buyer. In Raymond Clemens, editor, *The Voynich Manuscript*. Yale University Press, New Haven, 2016.

Kevin Knight and Kenji Yamada. A computational approach to deciphering unknown scripts. In *Proceedings of the ACL Workshop on Unsupervised Learning in Natural Language Processing*, pages 37–44, 1999.

Gabriel Landini. Evidence of linguistic structure in the Voynich manuscript using spectral analysis. *Cryptologia*, 25(4):275–295, 2001.

Gabriel Landini and René Zandbergen. The European Voynich Alphabet (EVA). https://www.voynich.nu/transcr.html, 1998.

George Lasry. Deciphering historical syllabic ciphers. In *Proceedings of the 7th International Conference on Historical Cryptology (HistoCrypt 2024)*, volume 53 of *NEALT Proceedings Series*, 2024.

Colin Layfield and Lisa Fagin Davis. Singulion structure and the Voynich manuscript. *Digital Medievalist*, 19, 2026. doi: 10.4000/16k0a. Published 19 October 2026.

Wentian Li and Kunihiko Kaneko. Long-range correlation and partial $1/f^{\alpha}$ spectrum in a noncoding DNA sequence. *Europhysics Letters*, 17(7):655–660, 1992.

Luke Lindemann and Claire Bowern. Character entropy in modern and historical texts: comparison metrics for an undeciphered manuscript. arXiv:2010.14697, 2021.

Beáta Megyesi, Crina Túdor, Benedek Láng, and Anna Lehofer. Key design in the early modern era in Europe. In *Proceedings of the 4th International Conference on Historical Cryptology (HistoCrypt 2021)*, pages 121–130, 2021. doi: 10.3384/ecp183165.

Beáta Megyesi, Crina Túdor, Benedek Láng, Anna Lehofer, Nils Kopal, and Michelle Waldispühl. What was encoded in historical cipher keys in the early modern era? In *Proceedings of the 5th International Conference on Historical Cryptology (HistoCrypt 2022)*, pages 159–167, 2022. doi: 10.3384/ecp188404.

Aloys Meister. *Die Geheimschrift im Dienste der päpstlichen Kurie von ihren Anfängen bis zum Ende des XVI. Jahrhunderts*. Ferdinand Schöningh, Paderborn, 1906.

Marcelo A. Montemurro and Damián H. Zanette. Keywords and co-occurrence patterns in the Voynich manuscript: an information-theoretic analysis. *PLoS ONE*, 8(6):e66344, 2013.

Hugh O'Neill. Botanical observations on the Voynich MS. *Speculum*, 19(1):126, 1944.

David Oranchak, Sam Blake, and Jarl Van Eycke. The solution of the Zodiac Killer's 340-character cipher. arXiv:2403.17350, 2024. URL https://arxiv.org/abs/2403.17350.

Nicholas Pelling. *The Curse of the Voynich: The Secret History of the World's Most Mysterious Manuscript*. Compelling Press, Surbiton, 2006.

Nick Pelling. Fifteenth century cryptography. Cipher Mysteries, https://ciphermysteries.com, 2016. Lavinde 1379, Mantua 1401, Tranchedino c. 1450: table shapes, homophones, nulls, nomenclators.

Nick Pelling. Voynich zodiac Gemini and the workshop of Diebold Lauber. Cipher Mysteries, https://ciphermysteries.com/2018/09/10/voynich-zodiac-gemini-and-the-workshop-of-diebold-laub 2018. Reporting Koen Gheuens's 2016 comparison.

Marco Ponzi. The VM zodiac as a pictorial cycle: A comparative analysis. Guest post on Stephen Bax's site, https://stephenbax.net/?p=1755, 2016. 131 zodiac cycles scored on twenty features.

Sravana Reddy and Kevin Knight. What we know about the Voynich manuscript. In *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, pages 78–86, 2011. URL https://aclanthology.org/W11-1511/.

Gordon Rugg. An elegant hoax? A possible solution to the Voynich manuscript. *Cryptologia*, 28(1):31–46, 2004.

Gordon Rugg and Gavin Taylor. Hoaxing statistical features of the Voynich manuscript. *Cryptologia*, 41(3):247–268, 2017. doi: 10.1080/01611194.2016.1206753.

Andreas Schinner. The Voynich manuscript: evidence of the hoax hypothesis. *Cryptologia*, 31(2):95–107, 2007.

Claude E. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 28(4):656–715, 1949.

Cicco Simonetta. Regule ad extrahendum litteras zifferatas sine exemplo, 1474. Pavia, 4 July 1474; ed. P.-M. Perret, *Bibliothèque de l'École des chartes* 51 (1890).

Jorge Stolfi. A prefix-midfix-suffix decomposition of Voynichese words; the core–mantle–crust word paradigm. https://www.ic.unicamp.br/~stolfi/voynich/, 1997–2000.

John H. Tiltman. The Voynich manuscript: "the most mysterious manuscript in the world". *NSA Technical Journal*, 12(3):41–85, 1967. Reporting work begun in 1951; declassified.

Torsten Timm. How the Voynich manuscript was created. arXiv:1407.6639, 2014.

Torsten Timm and Andreas Schinner. A possible generating algorithm of the Voynich manuscript. *Cryptologia*, 44(1):1–19, 2020.

Juergen Wastl and Danielle Feger. VM408 folio 86v 'the rosette map': Elements of a mappa mundi and a map of the elements. figshare, https://figshare.com/articles/VM408_folio86v_The_Rosette_Map_Elements_of_a_mappa_mundi_and_a_map_of_the_Elements/903756, 2014.

René Zandbergen. The Voynich manuscript. https://www.voynich.nu, 1996–2026. Including the pages on the manuscript's description, the quire structure and page order, the illustrations, the transliteration files (ZL3b, 13 May 2025) and the IVTFF format definition. Accessed October 2026.

René Zandbergen. Earliest owners. In Raymond Clemens, editor, *The Voynich Manuscript*. Yale University Press, New Haven, 2016.

Paula Zyats, Erin Mysak, Jens Stenger, Marie-France Lemay, Anikó Bezur, and David D. Driscoll. Physical findings. In Raymond Clemens, editor, *The Voynich Manuscript*. Yale University Press, New Haven, 2016. Summarising the materials analysis, including that of McCrone Associates (2009).