THE VOYNICH MANUSCRIPT

Complete synthesis of the digital, cryptographic, geometric, biological, astronomical and audio research

Investigation of a possible coded or multi-modal structure in the manuscript

Eric GARCIA
Expert in Digital Research and AI

Research synthesis — September 2026

Working document: experimental results, hypotheses, limitations and validation paths.

Executive Summary

This research begins with a simple question: could the characters and illustrations of the Voynich Manuscript contain, in addition to their apparent textual form, a digital or informational structure that only becomes visible after transformation?

The study progressively explored several representations: text and glyphs, binary data, density, periodicity, audio signals, FSK, synchronization, ASCII, coordinates, DNA coding, ORFs, geometry, the golden ratio, spirals, astronomy, star groups, graphic comparison with modern “Light Language” systems, and a conceptual comparison with the Wow! Signal.

The main result is not a decoding. Several transformations produce regularities and patterns, but none currently provides a clear, reproducible and statistically validated message. The strongest next step is a controlled statistical protocol using a standardized Voynich corpus, systematic permutations and random control corpora preserving the same distributions.

> **Guiding principle: systematically distinguish what is observed, what is interpreted, and what remains hypothetical.**

1. Corpus and Starting Material

The reference document used in this research is the PDF copy of the manuscript supplied for the study. The copy contains 209 pages. Automatic text extraction from the PDF is very limited because the pages are primarily scans; visual analysis is therefore essential.

Folios f115r/f115v were used as a privileged experimental area. In the studied copy, they correspond to pages 115 and 116 and show long blocks of text accompanied by marginal stars.

The research was subsequently extended to other folios, especially astronomical pages and folio f68r3, as well as global or semi-global processing of the corpus.

1.1. Why f115r/f115v?

These pages are particularly useful for a digital experiment because they contain a large amount of text, few central illustrations, and clearly visible marginal elements. They therefore allow the structure of the text to be tested without the large illustrations dominating the analysis.

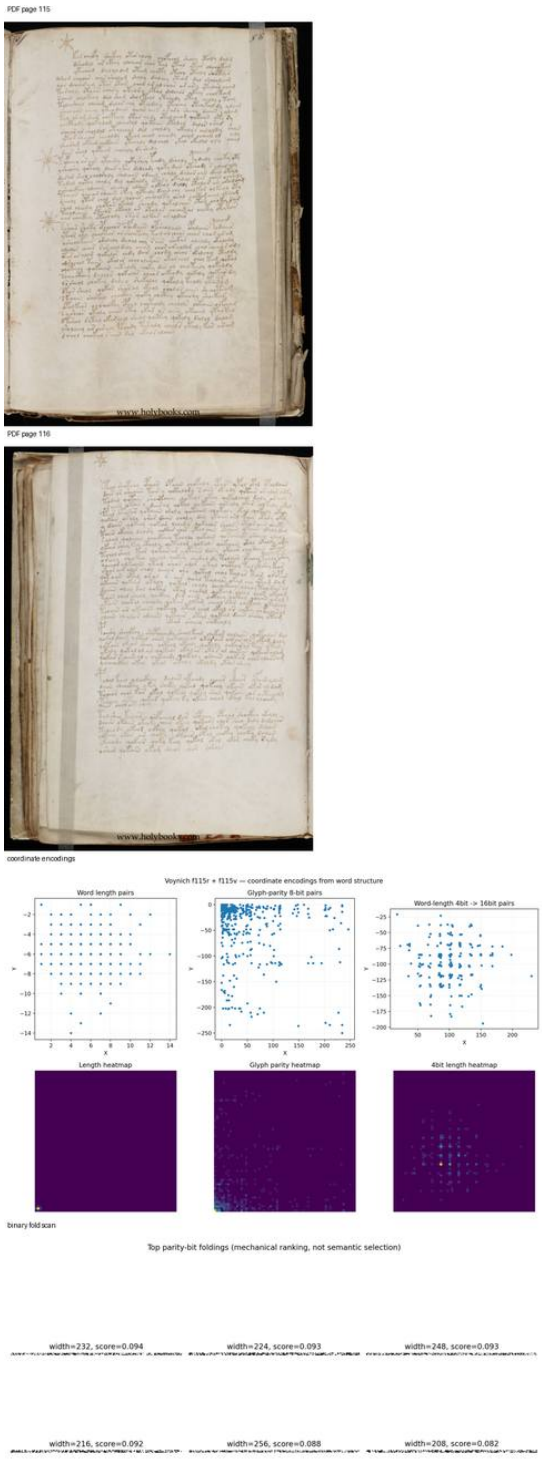

Figure 1 — Experimental montage resulting from transformations applied to f115/f116.

2. General Methodology

The approach consisted of successively transforming the same information in order to look for properties that would remain present despite a change of representation.

> **VOYNICH → glyphs → numerical values → bits → image / signal → candidate decodings**

Each transformation can nevertheless create artificial patterns. A visually attractive form is therefore not considered proof. The research distinguishes exploratory results from established conclusions.

2.1. Test families

• Textual: frequencies, word lengths, repetitions and positions.

• Binary: conversion to bits, reshaping, inversions, offsets and density.

• Audio: conversion of bits into FSK and spectral analysis.

• ASCII: grouping into bytes, offsets, inversion and text searches.

• Biological: conversion into four states A/C/G/T, 24 mappings and ORF searches.

• Geometric: coordinates, symmetries, circles, spirals and the golden ratio.

• Astronomical: stars, the Pleiades, celestial configurations and date searches.

• Comparative: cautious comparison with “Light Language” and reflection on the Wow! Signal.

3. Binary Transformation

The first axis considers the manuscript as a sequence of symbols that can be converted into numbers and then into bits. Several conventions were tested because there is no independent reason to choose one unique character-to-number mapping.

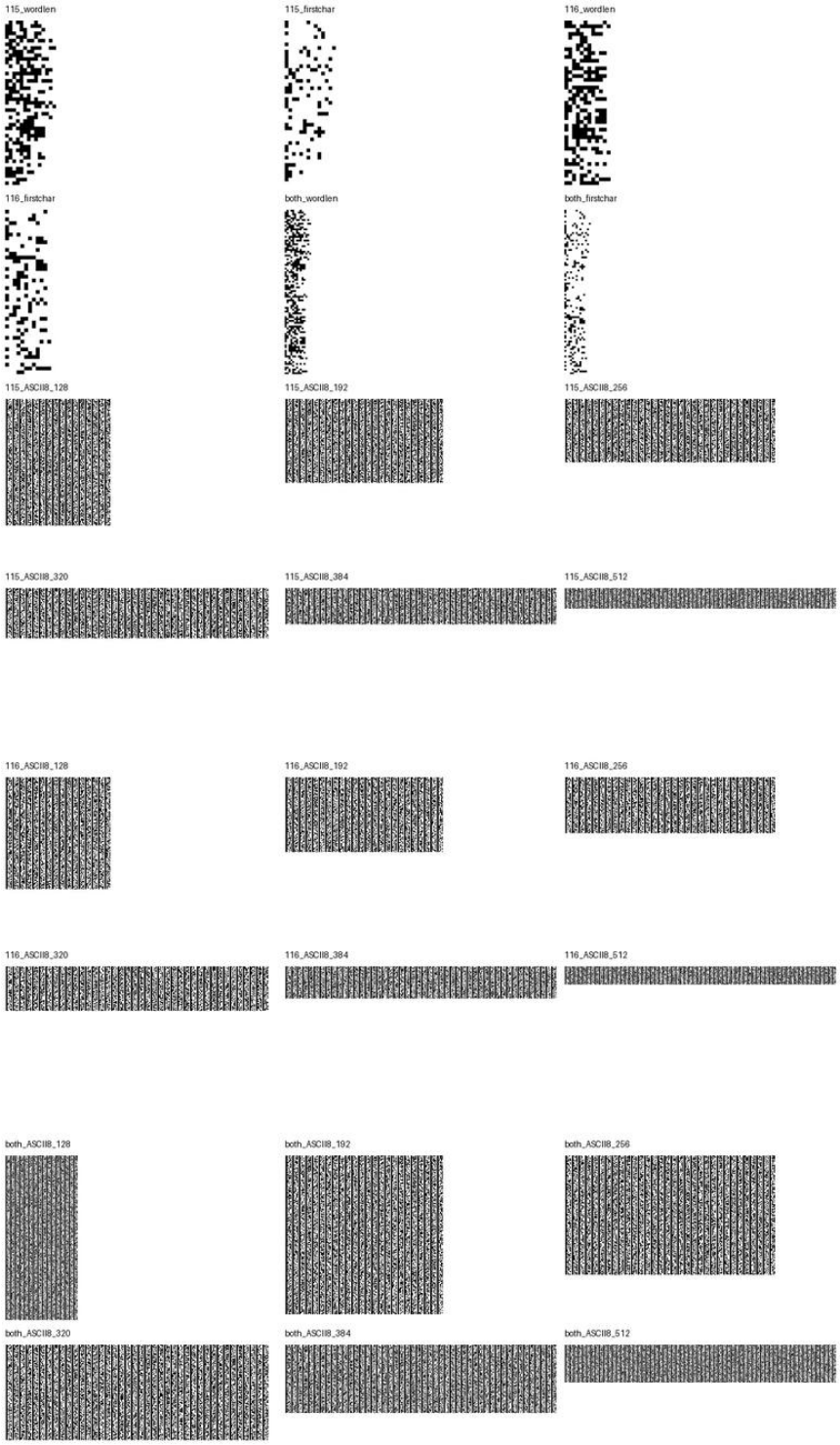

Figure 2 — Binary representation of the studied textual material.

Figure 3 — Reshaping the binary sequence at different widths.

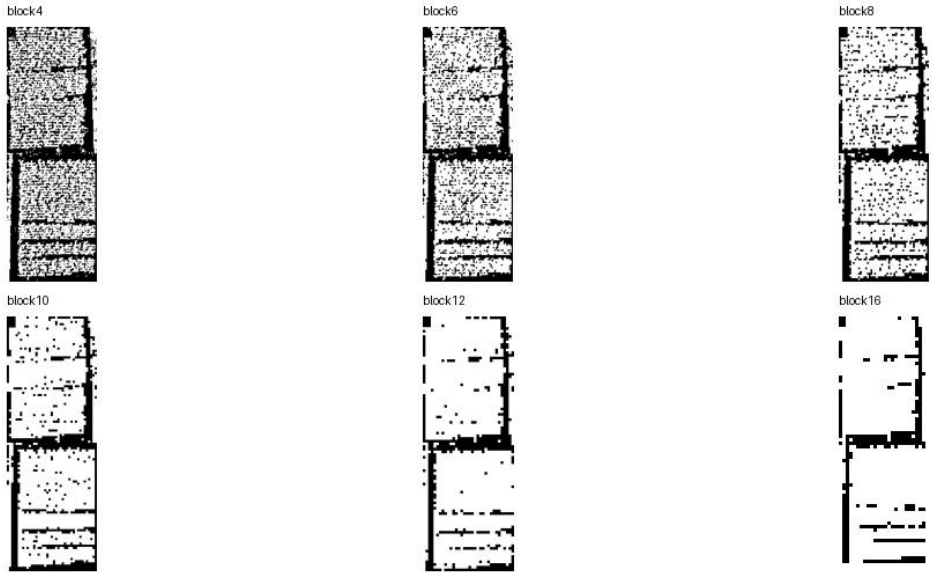

Figure 4 — Density analysis of binary states.

Structures appear: bands, repetitions and density regions. However, these patterns are compatible with page layout, glyph repetition and the chosen representation. No obvious hidden message was recovered.

4. Search for Periodicities

The binary sequence was examined as a discrete series to look for repetitions or preferred frequencies. A broader analysis was also carried out across manuscript pages.

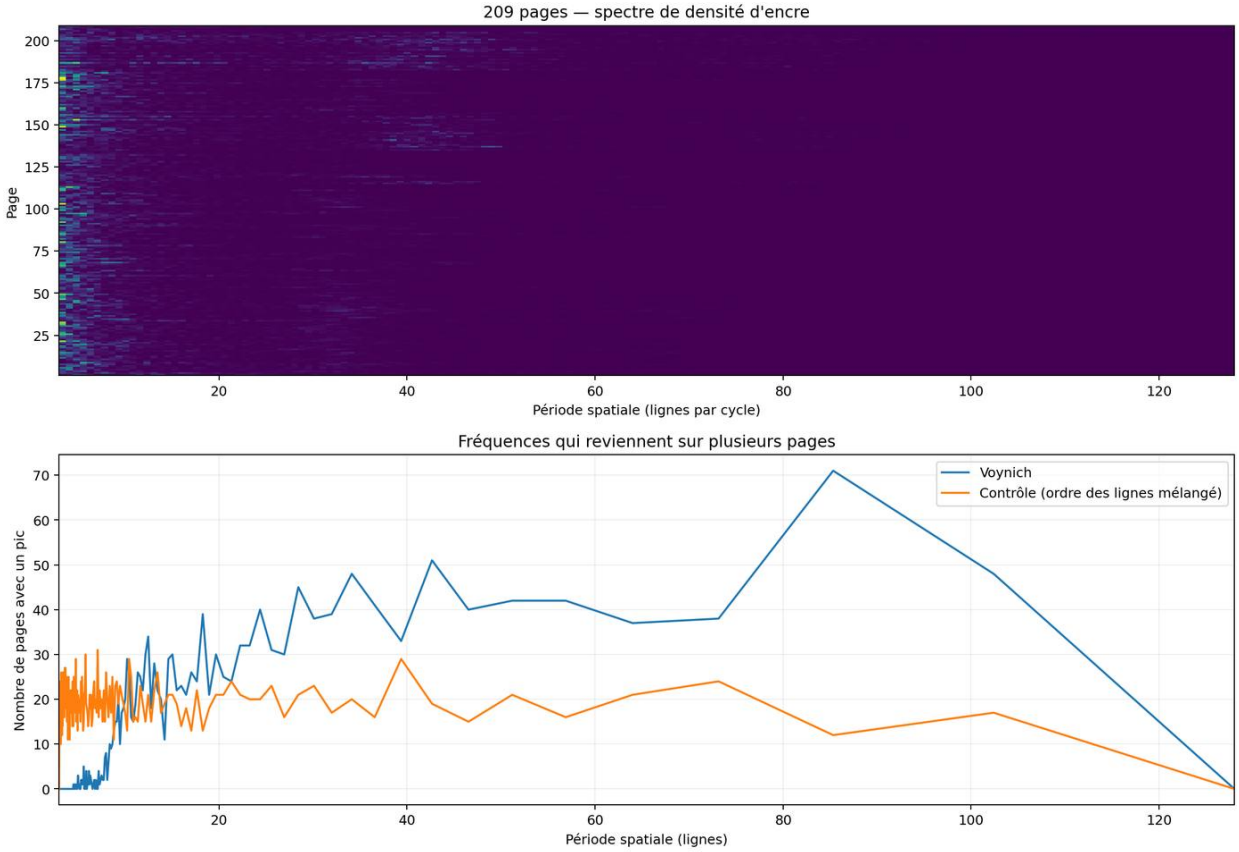

Figure 5 — Scan of periodicities and recurrent structures across the corpus.

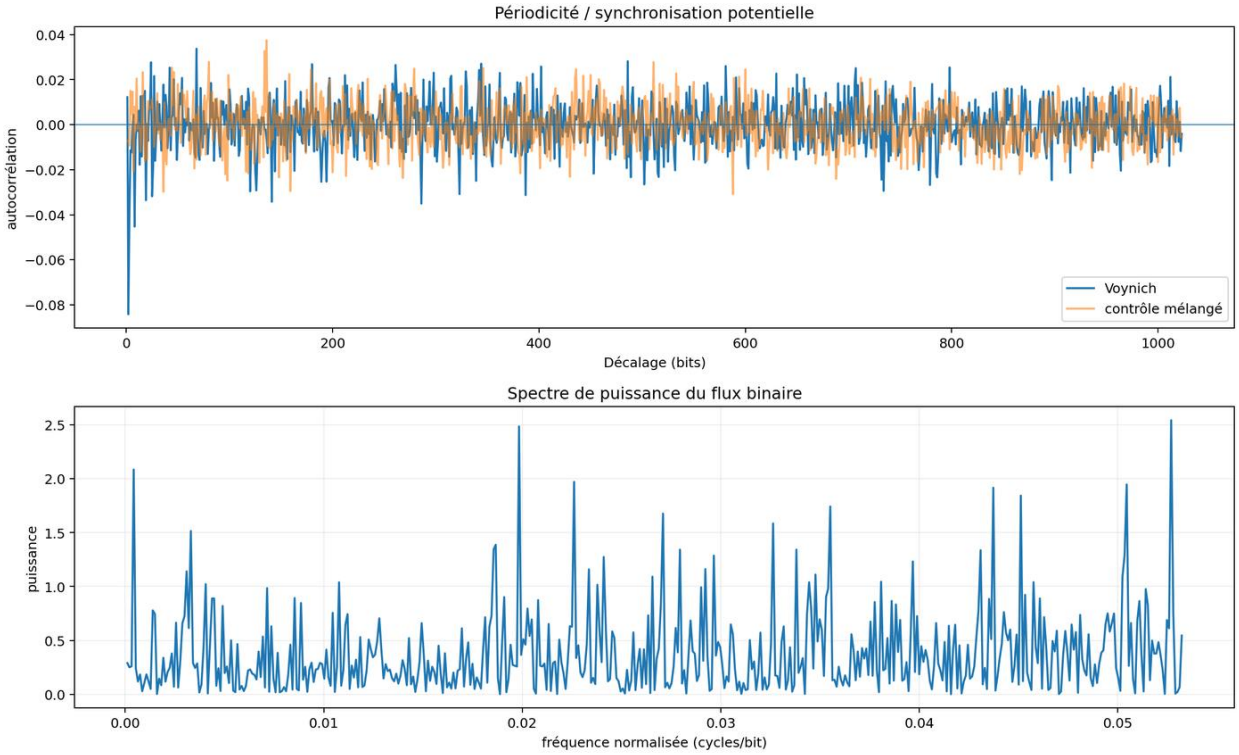

Figure 6 — Periodicity experiment using a signal-like representation.

The result is strongly negative in the narrow sense: no single periodicity emerges as an indisputable signature of an external encoding. The experiment nevertheless supports examining the manuscript as a data sequence rather than only as a text.

5. Geometric and Spatial Research

Words and characters were converted into positions, lengths and values to test whether the spatial organization could generate a figure when the presumed meaning of the glyphs is removed.

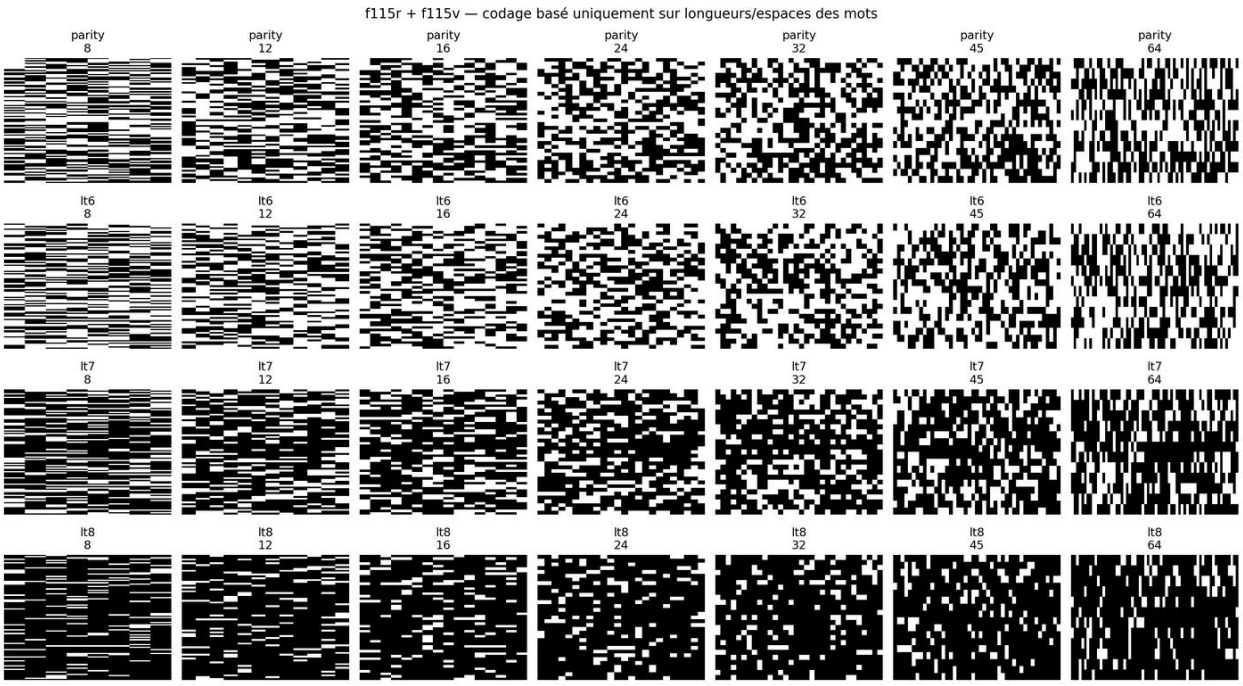

Figure 7 — Rasterization based on word lengths.

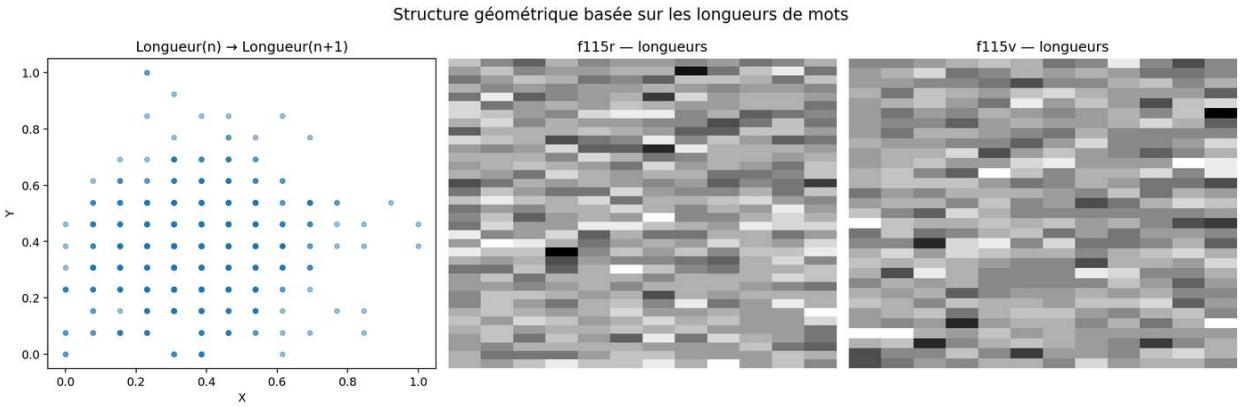

Figure 8 — Spatial representation of words and their lengths.

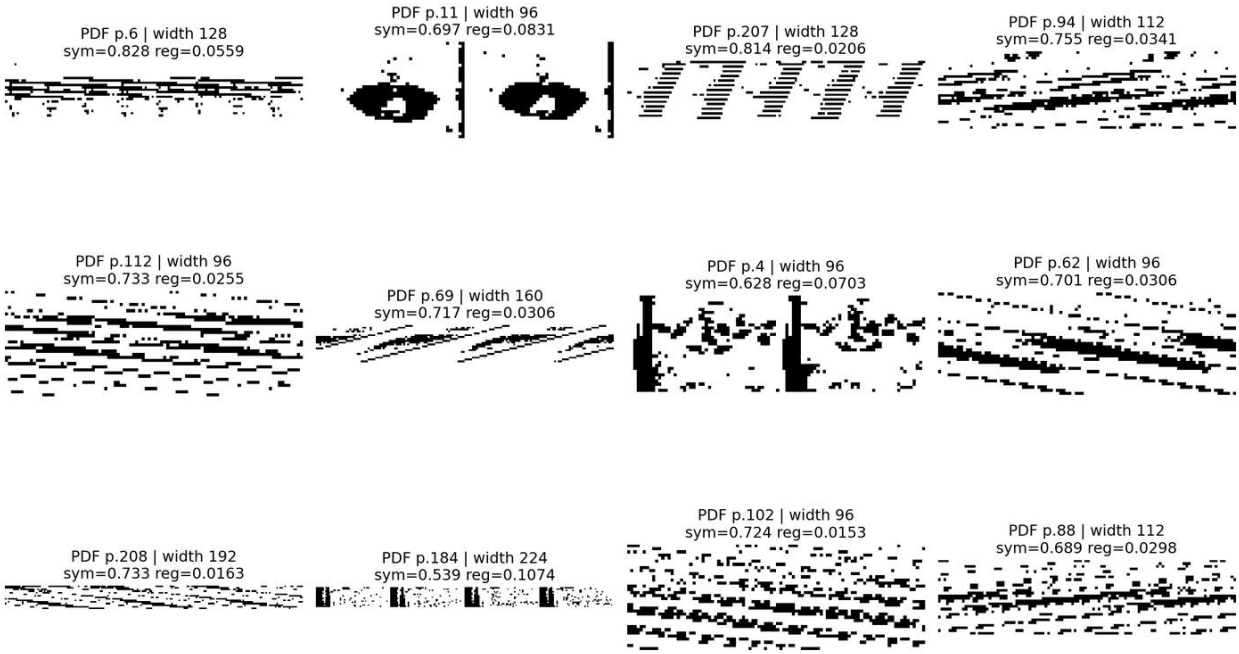

Figure 9 — Geometric scan of a page sample using several reshaping widths.

About fifty pages were subjected to a search for regularity and symmetry. The highest scores remain mainly explainable by lines, margins and text blocks. No sufficiently stable hidden drawing was identified.

6. Coordinates Derived from Binary Data

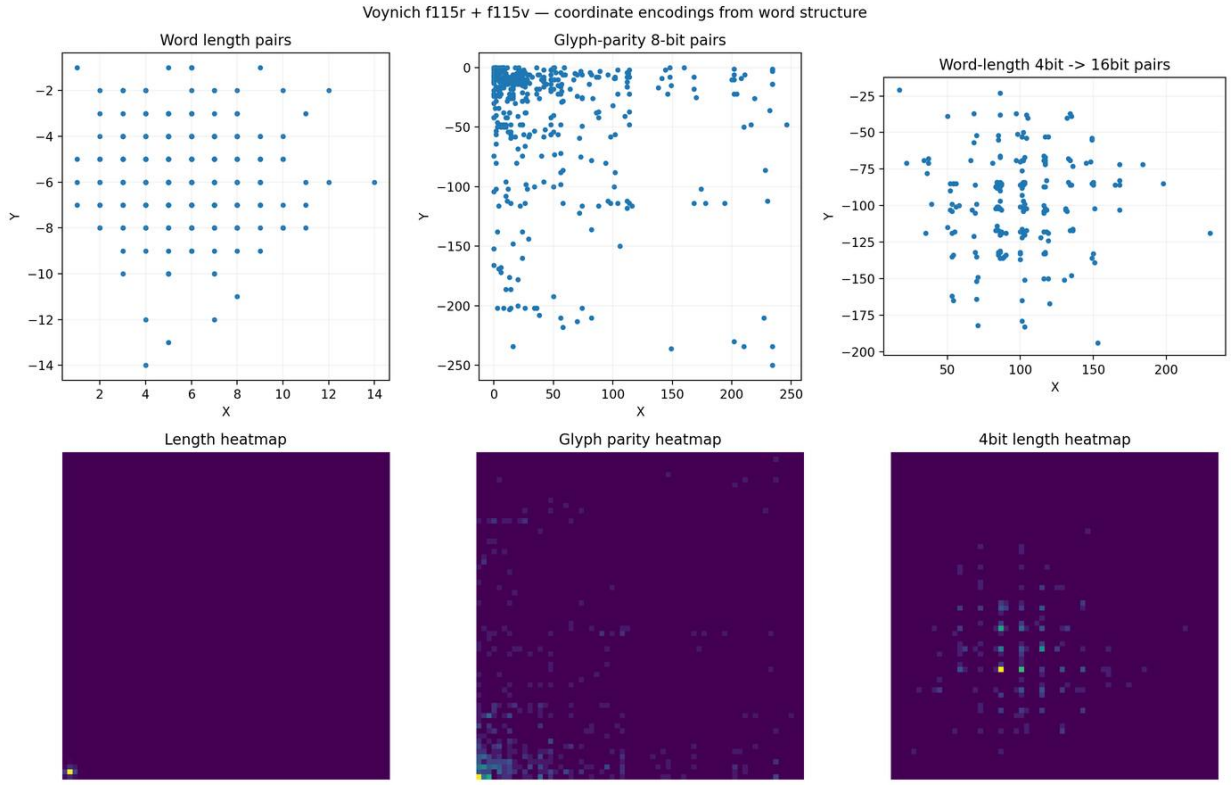

Figure 10 — Coordinates derived from f115/f116 data.

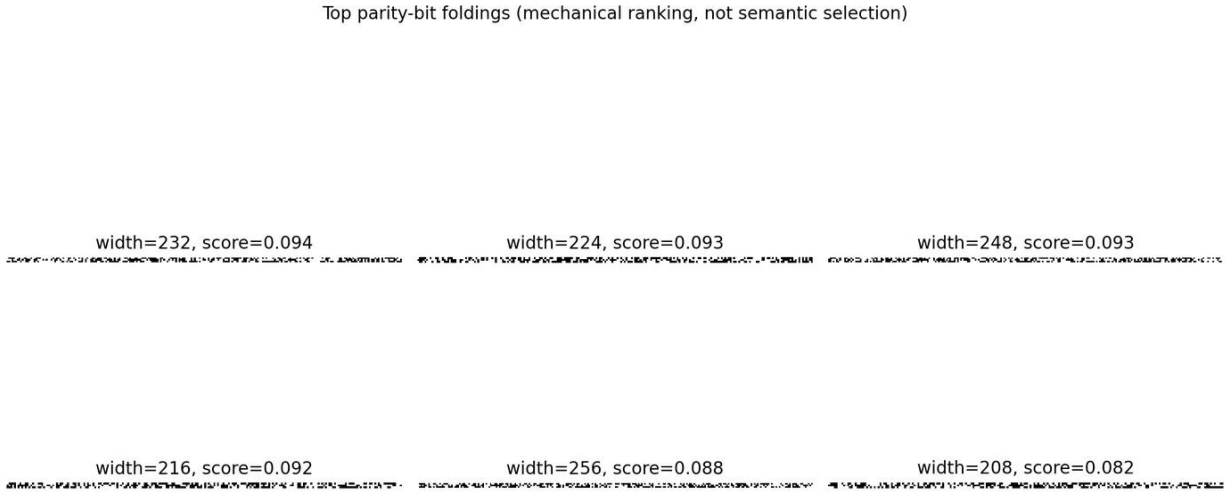

Figure 11 — Coordinate scan from binary groups.

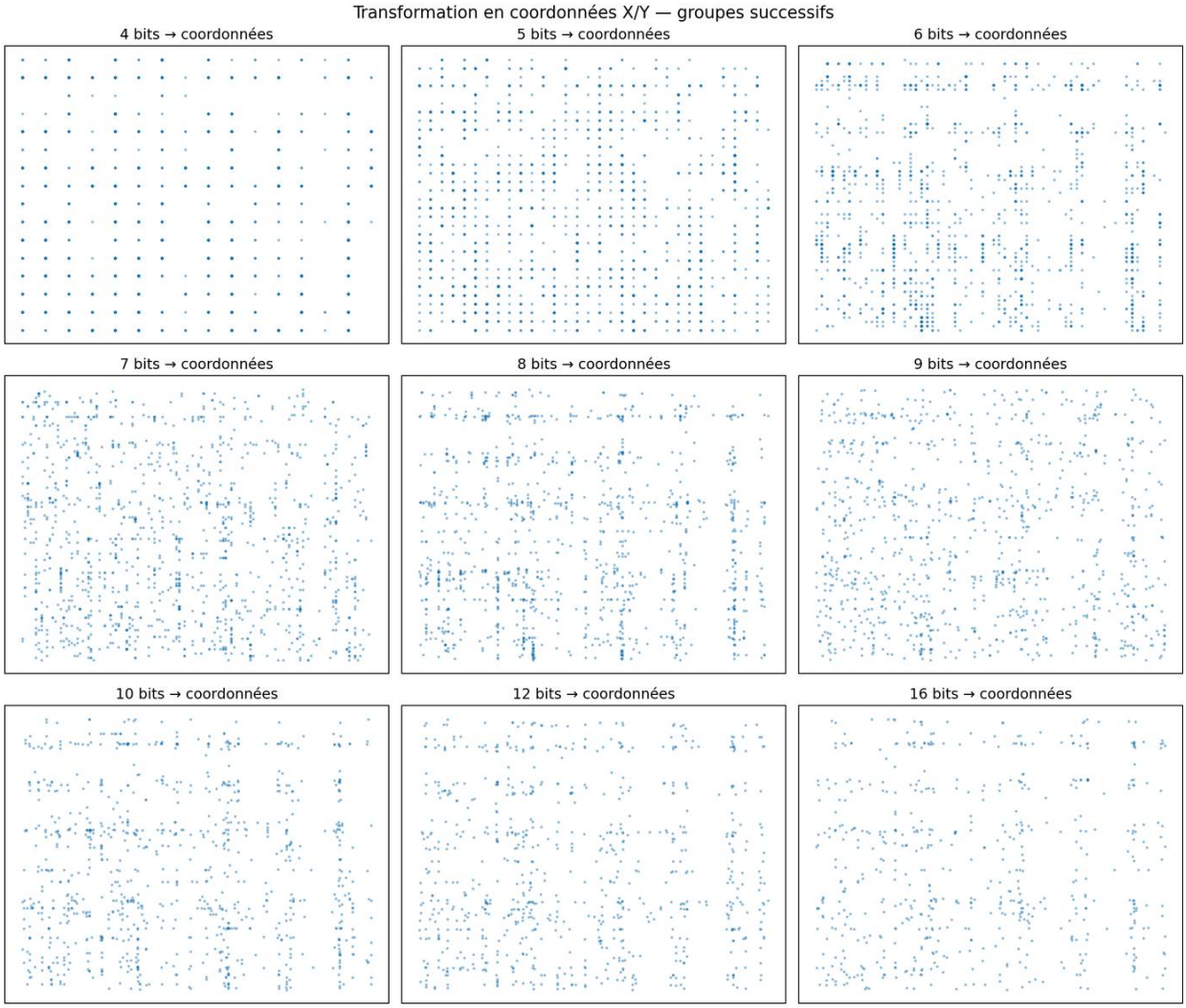

Figure 12 — Representation using bit groups and coordinates.

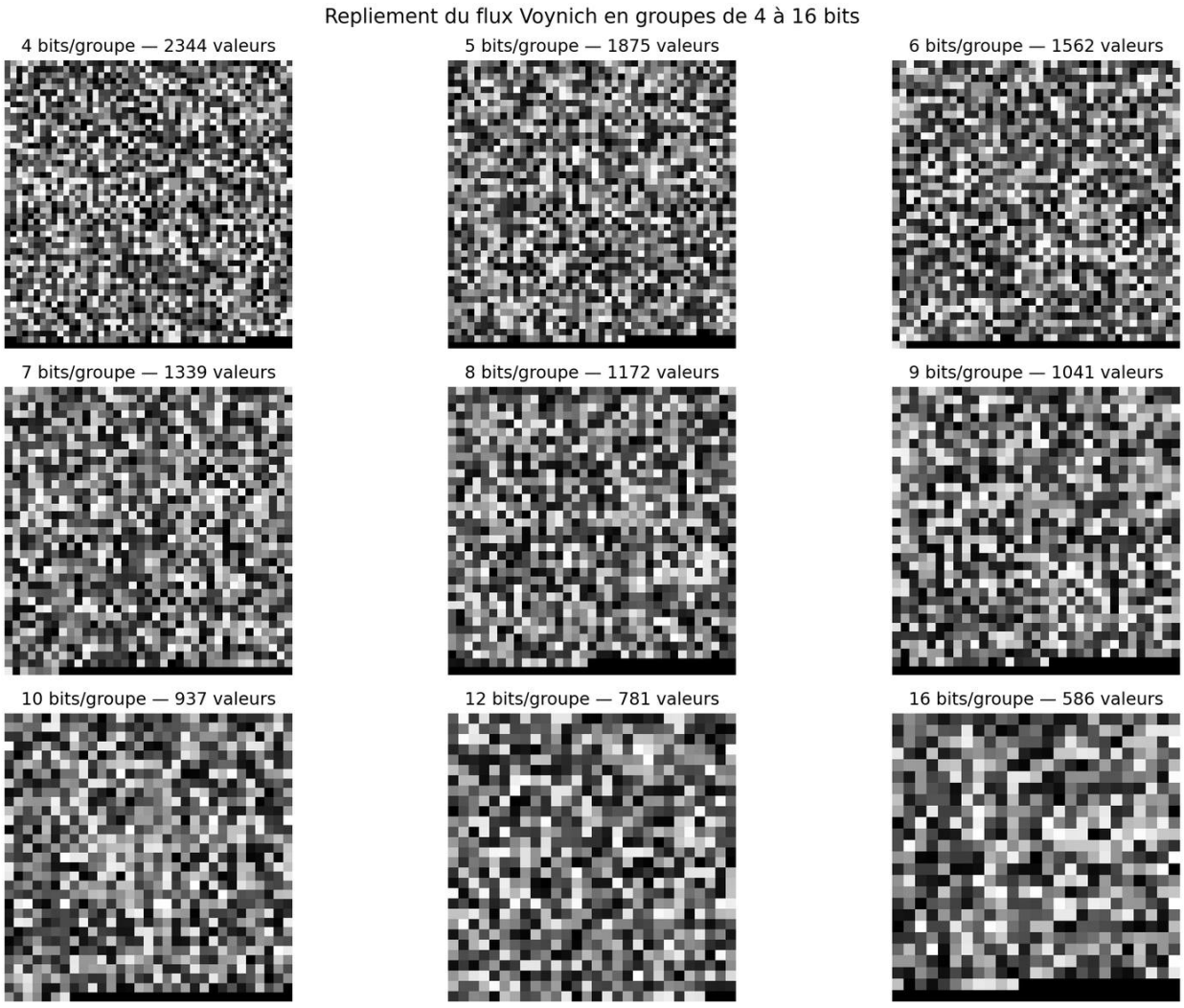

Figure 13 — Image reshaping of bit groups.

The resulting point clouds do not produce a clearly recognizable map, constellation or spiral. The main value of this branch is methodological: it provides useful negative controls.

7. FSK Audio Experiment

A central idea in the research was to treat a binary sequence as temporal information. Two binary states can be represented by two distinct frequencies: this is the principle of FSK (Frequency Shift Keying).

An initial experimental construction used frequencies around 900 Hz and 1900 Hz. Several WAV/PCM and MP3 files were generated and examined using spectrograms.

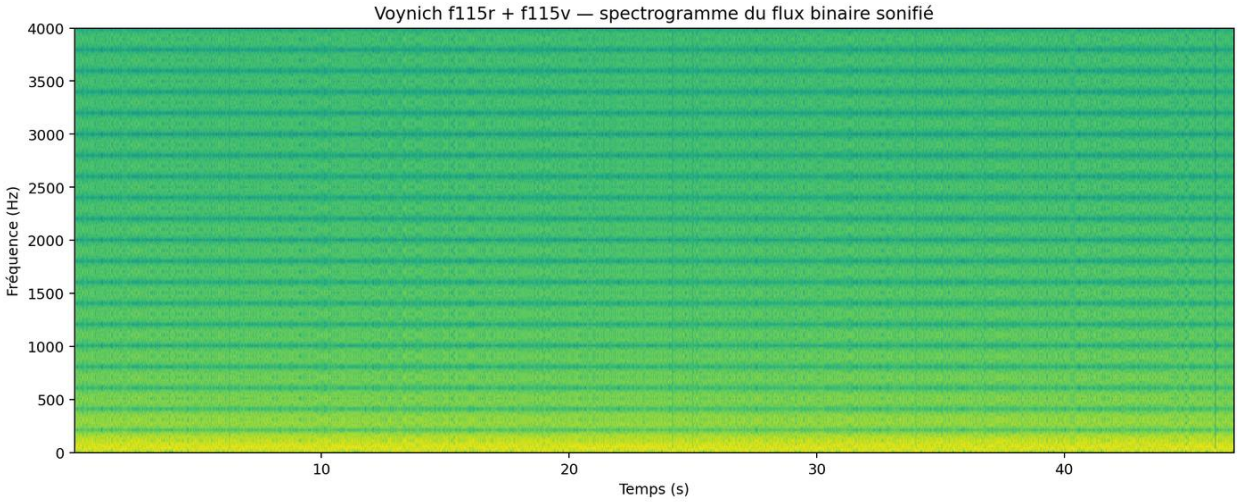

Figure 14 — Spectrogram of the audio signal derived from f115/f116.

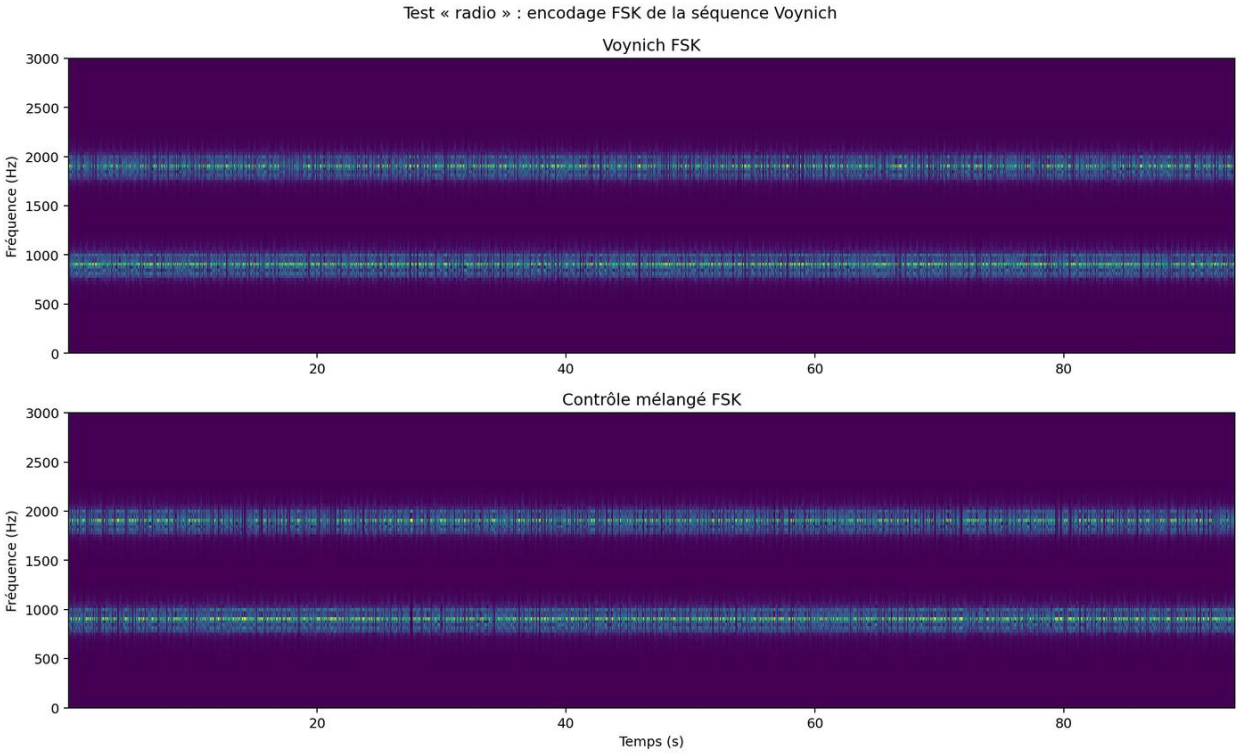

Figure 15 — Spectrogram of the FSK experiment.

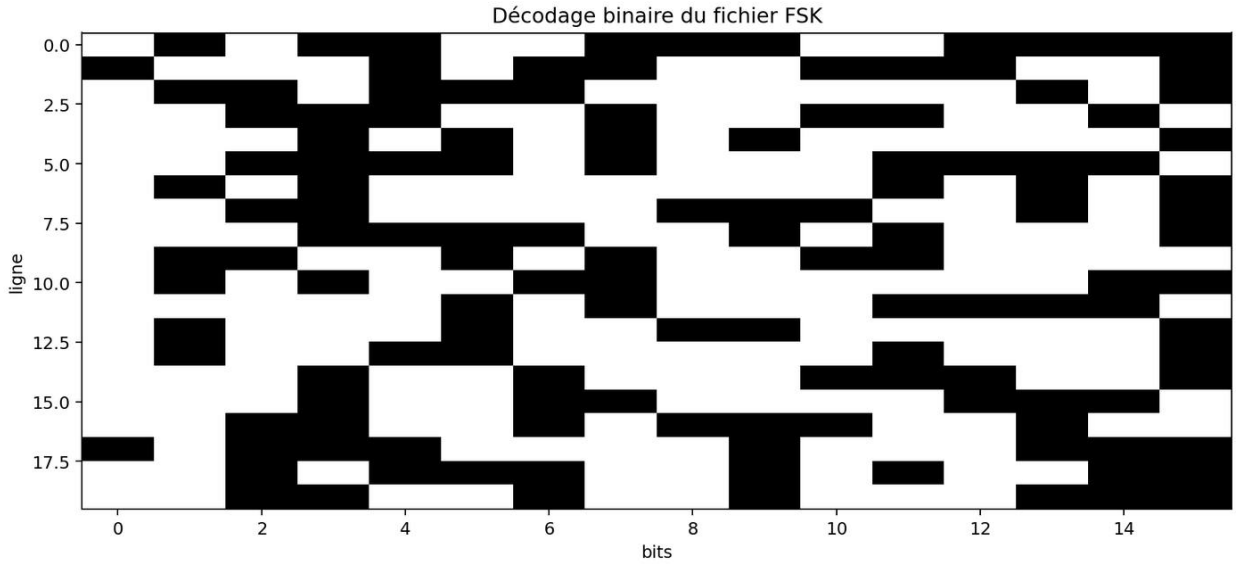

Figure 16 — Binary decoding representation of the FSK signal.

In the research notes, frequency values were reported around 902.5 Hz and 1892.5 Hz, with a difference close to 990 Hz and a rate of approximately 87.5 bits/s. These values describe the experiment as constructed and analyzed; they do not prove that a historical radio signal was present in the manuscript.

8. Synchronization and ASCII Search

After generating the signal, the bits were grouped into 8-bit bytes. Several offsets were tested, together with bit inversion, bit order reversal and several textual representations.

One experimental sequence obtained was:

Y..9n.92.A=.P.0..Qe0S...D.L..9..2..

This string contains printable ASCII characters, but it does not constitute a readable message. Broader searches were also performed using UTF-8, hexadecimal and Base64.

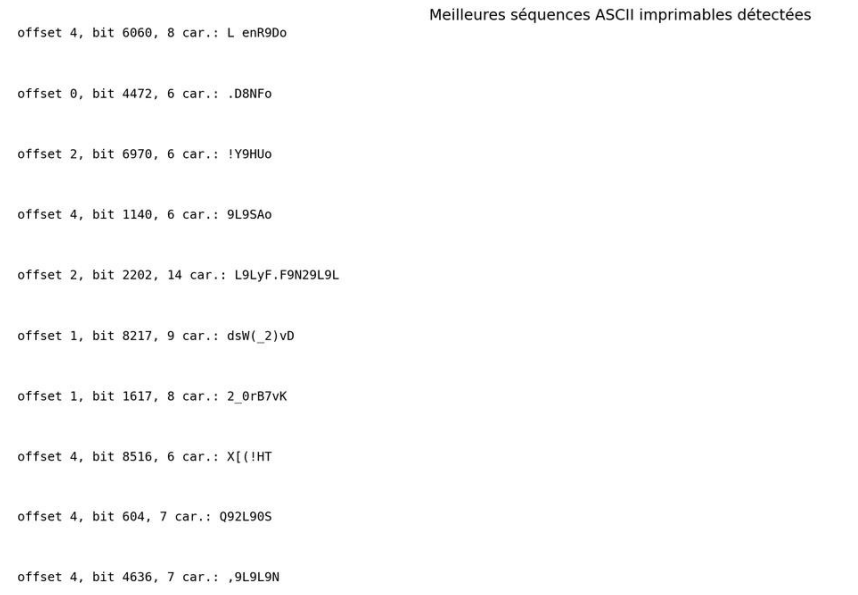

Figure 17 — Scan of ASCII outputs.

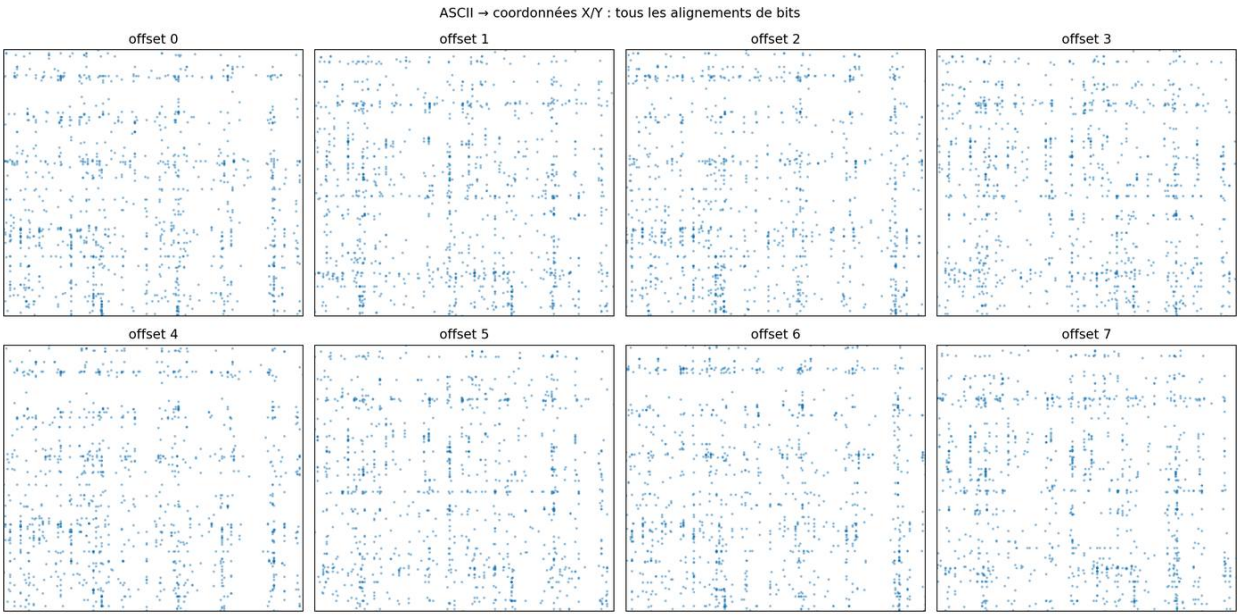

Figure 18 — ASCII coordinate tests across several offsets.

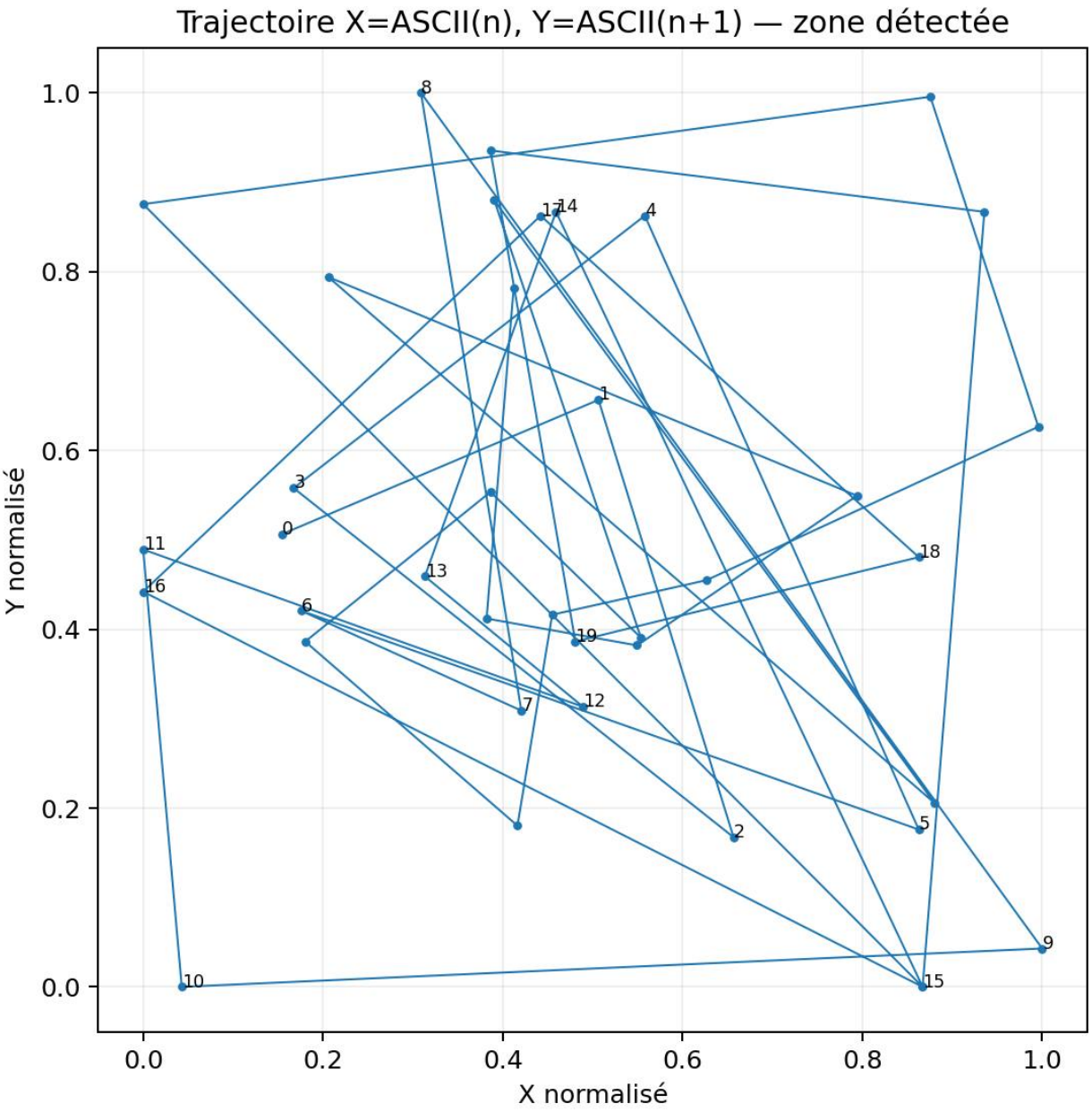

Figure 19 — Coordinate trajectory derived from ASCII values.

Conclusion: no stable linguistic string was obtained. The result is therefore exploratory rather than a decoding.

8.1. Methodological note: the “HELLO WORLD” control

During the broader research process, a separate binary control sequence was decoded as an ASCII example. The correctly encoded sequence produces “HELLO WORLD!”. This was a control/example for the binary-to-ASCII conversion and was not extracted from the Voynich Manuscript. The Voynich-derived FSK output above did not produce that message.

9. DNA Hypothesis

The next idea was to reduce the information to four states, allowing a natural correspondence with the four DNA bases: A, C, G and T. Mathematically, four states correspond to two bits.

With four symbols there are 4! = 24 possible permutations. The different correspondences were compared. One particular mapping, noted GACT in the experiment, stood out on the studied data.

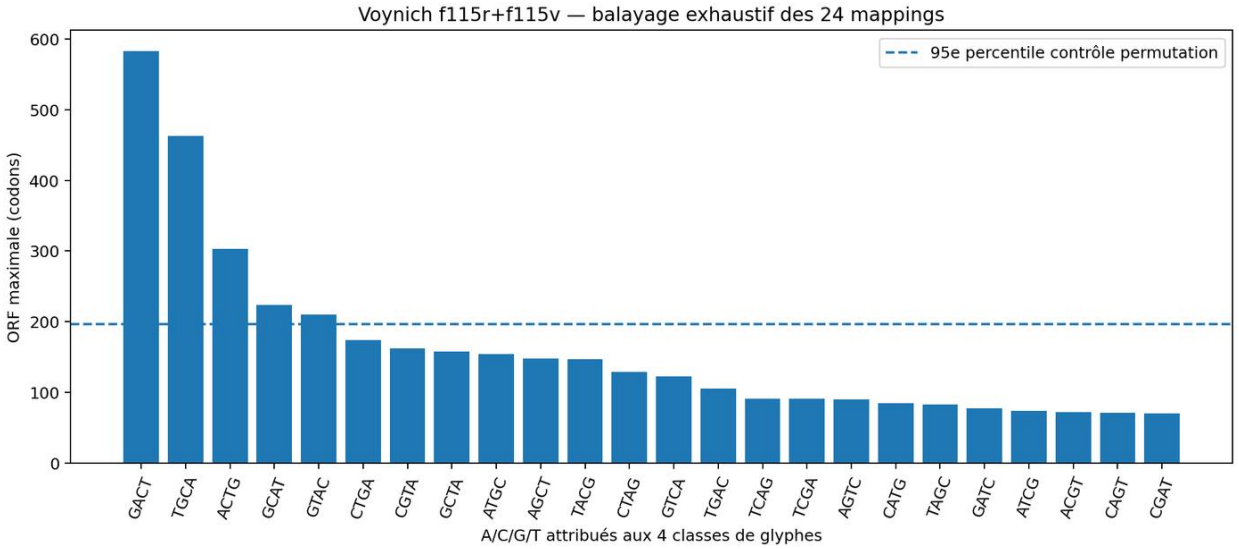

Figure 20 — Comparison of the 24 possible mappings to A/C/G/T.

Meilleur mapping (ORF max = 583 codons, ATG=23, stops=25, GC=53.3%)

AAC ATG TTC GCT GTC AAC TCG GAA GTC AGT GCG AAA AGG AGT TAG ATC AAG GGA AAG TGG TCA AGA TCA AGA ACA AGT GGA AGT GGG TGG AAG ATC AGA GTG GAA GTG GAG AGA GCG TCA

Contrôle: 200 permutations du texte par mapping; médiane ORF=136, 95e percentile=196 codons.

Figure 21 — Result of the exploratory GACT mapping.

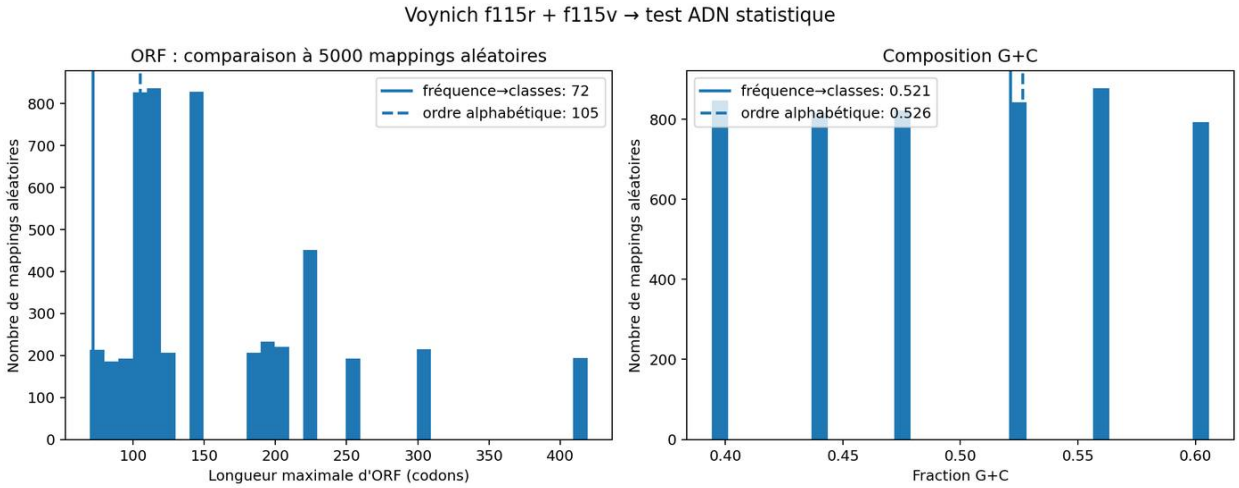

Figure 22 — Statistics for the exploratory DNA result.

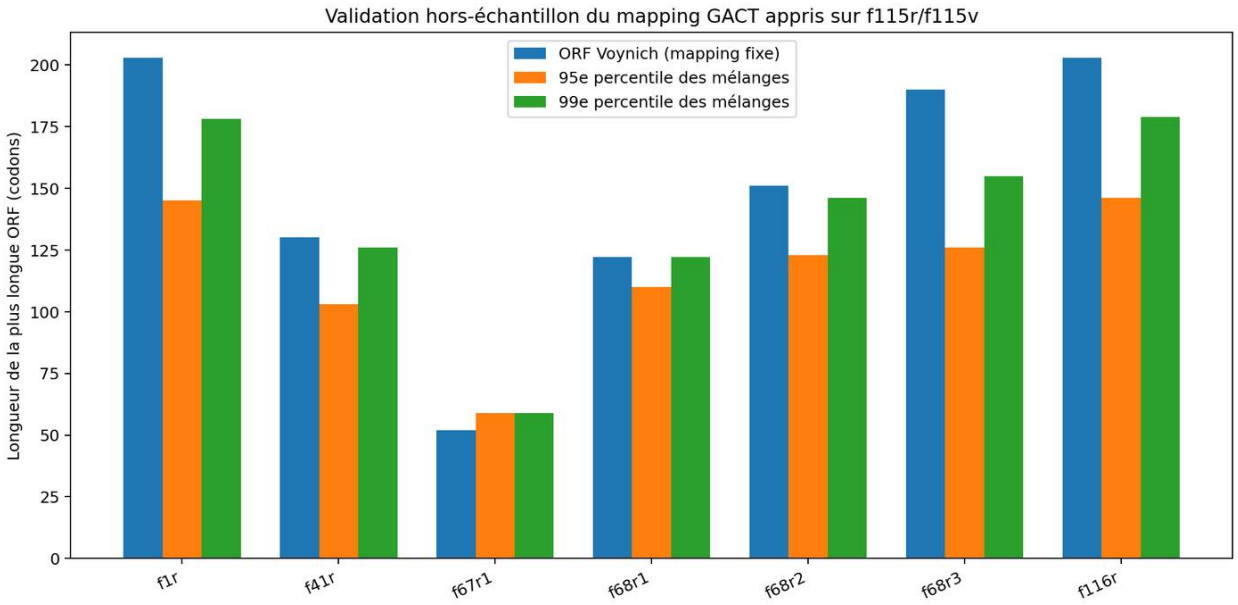

Figure 23 — Attempted independent validation of the DNA signal.

An obtained excerpt does indeed resemble a biological sequence, for example: GTCTCTTCAAGTGATATCCACATCCTCGCGGTCCAATGGATGGGCCAAT... However, such a resemblance is expected whenever an arbitrary sequence is converted into four letters.

10. ORF Search and Biological Limitations

The DNA sequence was examined for motifs such as ATG, reading frames and potentially long sequences. Comparisons with shuffled sequences were also used.

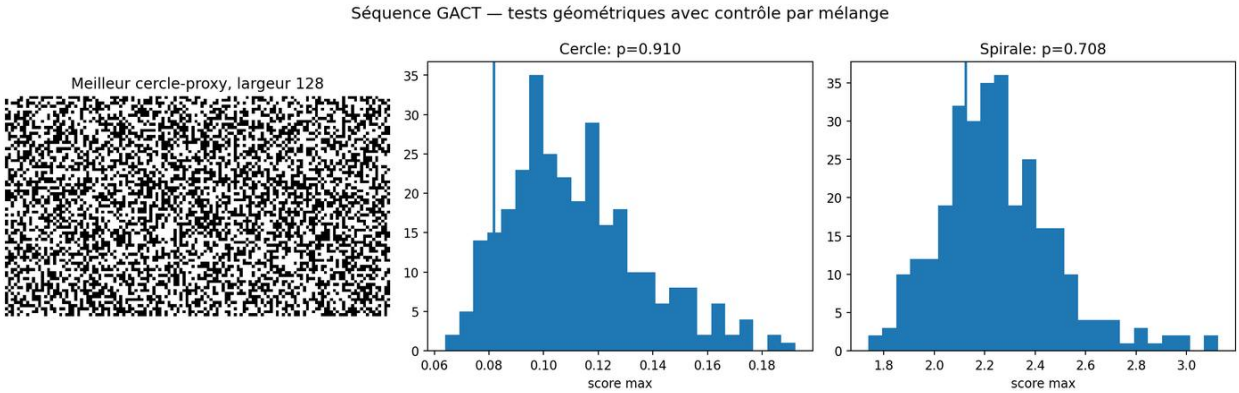

Figure 24 — Complementary test of the GACT mapping in a structured representation.

The critical point is that an ORF can occur by chance. For a biological hypothesis to become convincing, the observed signatures would need to clearly exceed random-sequence distributions and reproduce across independent sections.

The DNA branch therefore remains an interesting experimental hypothesis, but it does not currently support a conclusion that a genome or real biological information is present.

11. Golden Ratio and Spirals

Another line of inquiry looked for a proportion close to the golden ratio φ ≈ 1.618 in the manuscript's circular diagrams.

Figure 25 — Circular area studied for proportions.

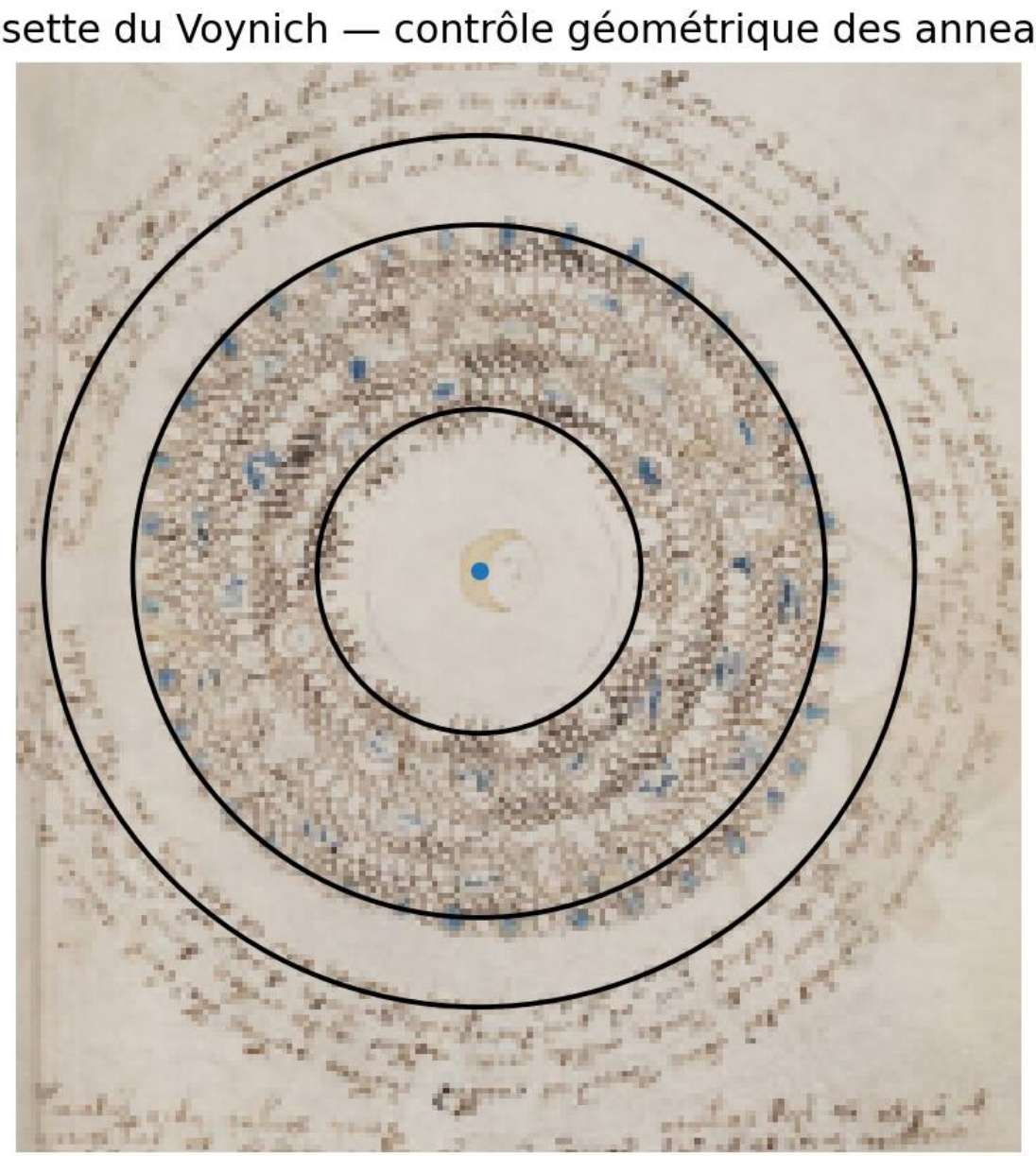

Figure 26 — Radius measurements and tested ratios.

Figure 27 — Logarithmic spiral fit on the studied motif.

Three prominent radii measured in one experiment were approximately 29 px, 62 px and 78 px, giving ratios of about 2.138, 1.258 and 2.690. They do not constitute a golden-ratio signature.

The logarithmic fit performed on a blue motif was approximately $r = 227.8 \exp(-0.0122\theta)$, with a factor per turn close to 0.926. The behavior is clearly different from a classical golden-growth spiral.

Conclusion: the circular geometry is real and interesting, but the golden ratio was not demonstrated by these measurements.

12. Astronomical Research

The astronomical branch is qualitatively different from the digital transformations: it relies on elements explicitly drawn in the manuscript. The celestial folios contain stars, circles and organizations that can be compared with configurations of the sky.

Figure 28 — General view of the studied f68r3 folio.

Figure 29 — Detail of the celestial diagram.

Figure 30 — Experimental detection/segmentation of stellar elements.

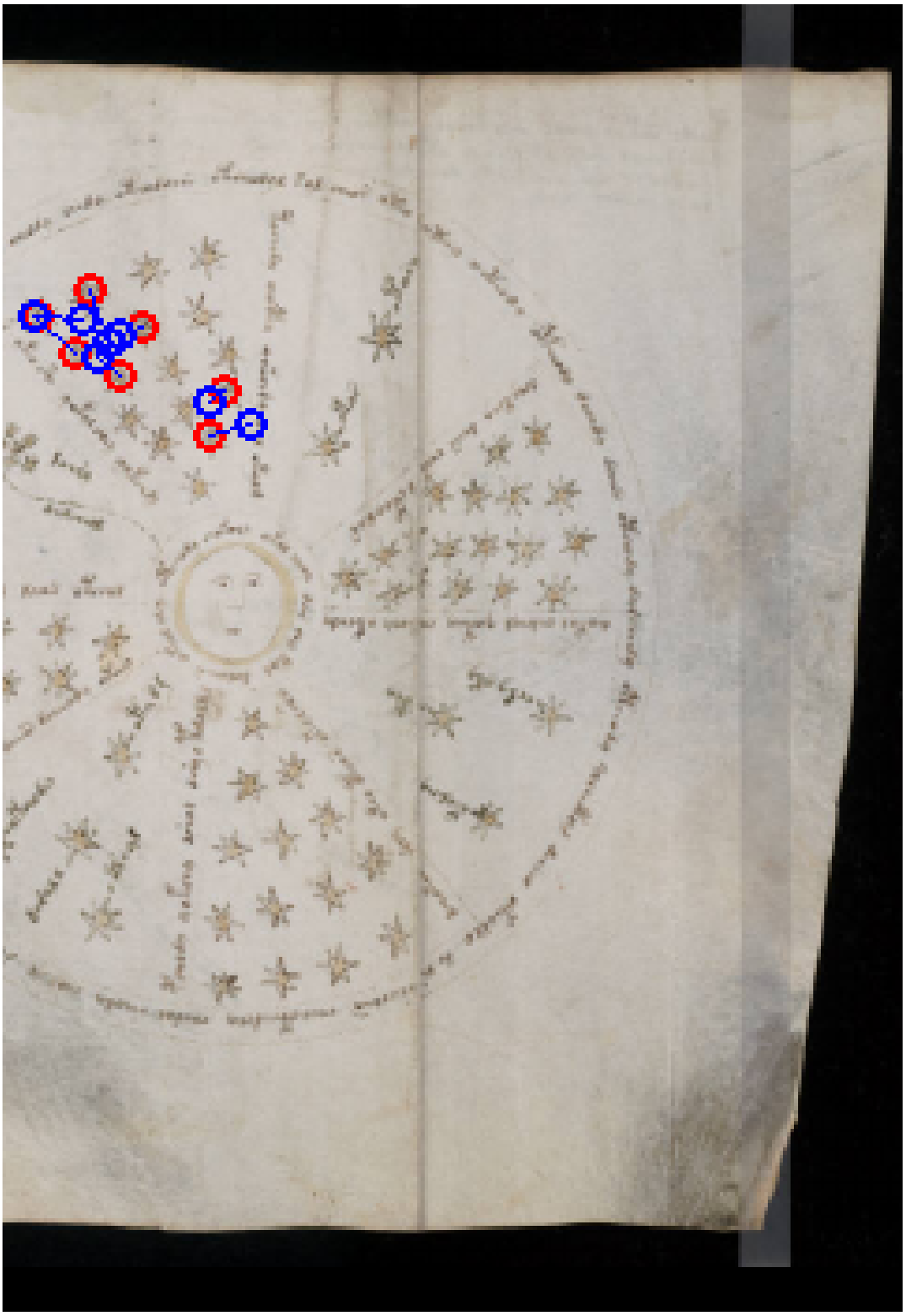

Figure 31 — Exploratory fit around the seven-star group traditionally associated with the Pleiades.

The seven-star group on f68r3 is a particularly interesting starting point for astronomical analysis. Precise identification of every star and dating of the drawing remain separate questions.

Exploratory astronomical calculations showed that there are historical dates on which the Moon can lie very close to the Pleiades region. Such a coincidence is interesting, but it is not sufficient to date the manuscript: several celestial objects must be compared simultaneously.

13. Comparison with “Light Language”

The comparison arose from a visual resemblance between some Voynich glyphs and modern alphabets presented under the name “Light Language”: loops, hooks, spirals and repetitive forms.

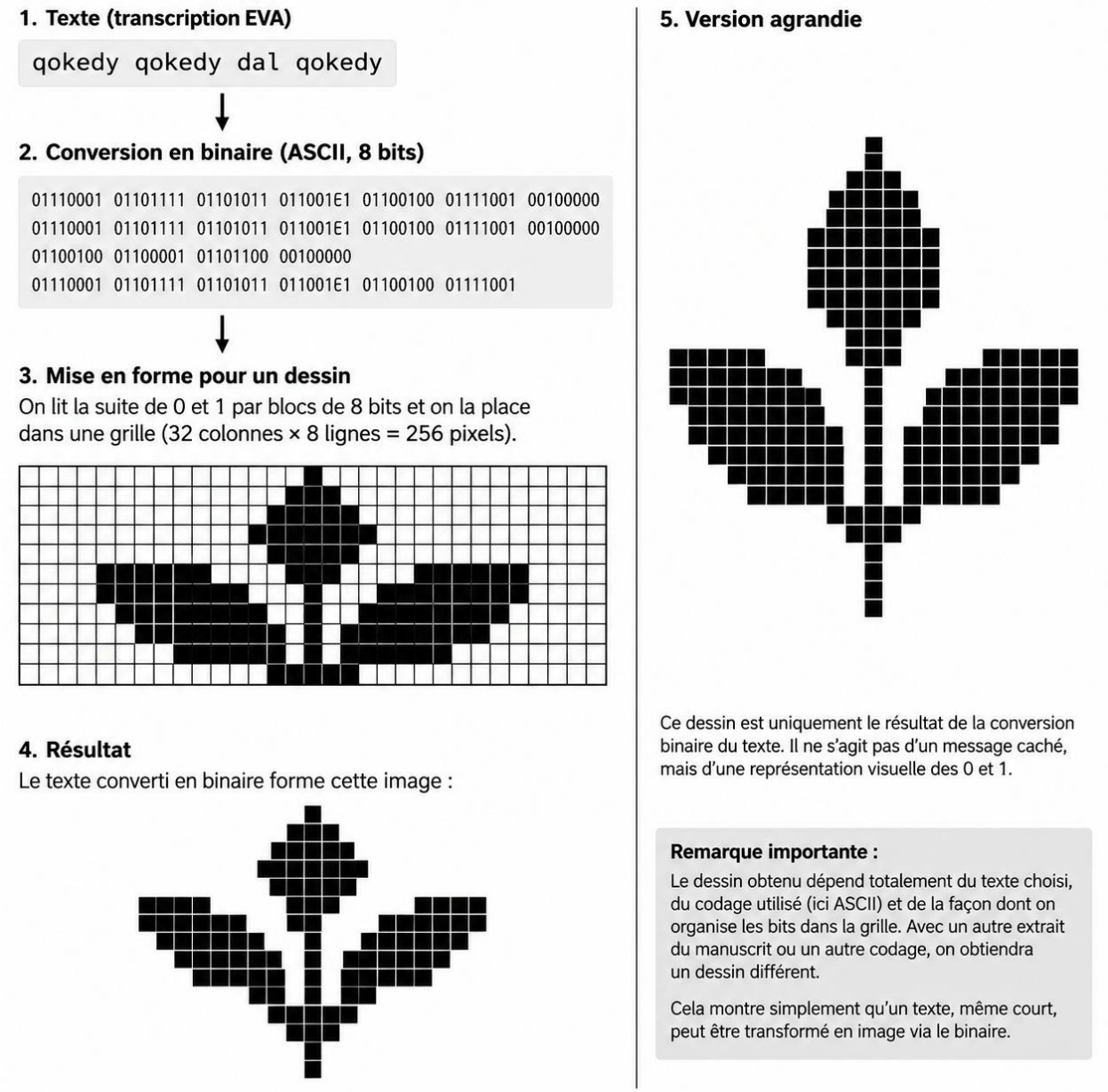

Figure 32 — Illustrative graphic used in the research dossier for the comparative discussion.

The resemblance remains graphical. No Voynich → Light Language correspondence table has been found that decodes the text using a reproducible grammar and vocabulary. This branch must therefore remain separate from any historical conclusion.

14. Hypothesis of a Scrambled Signal

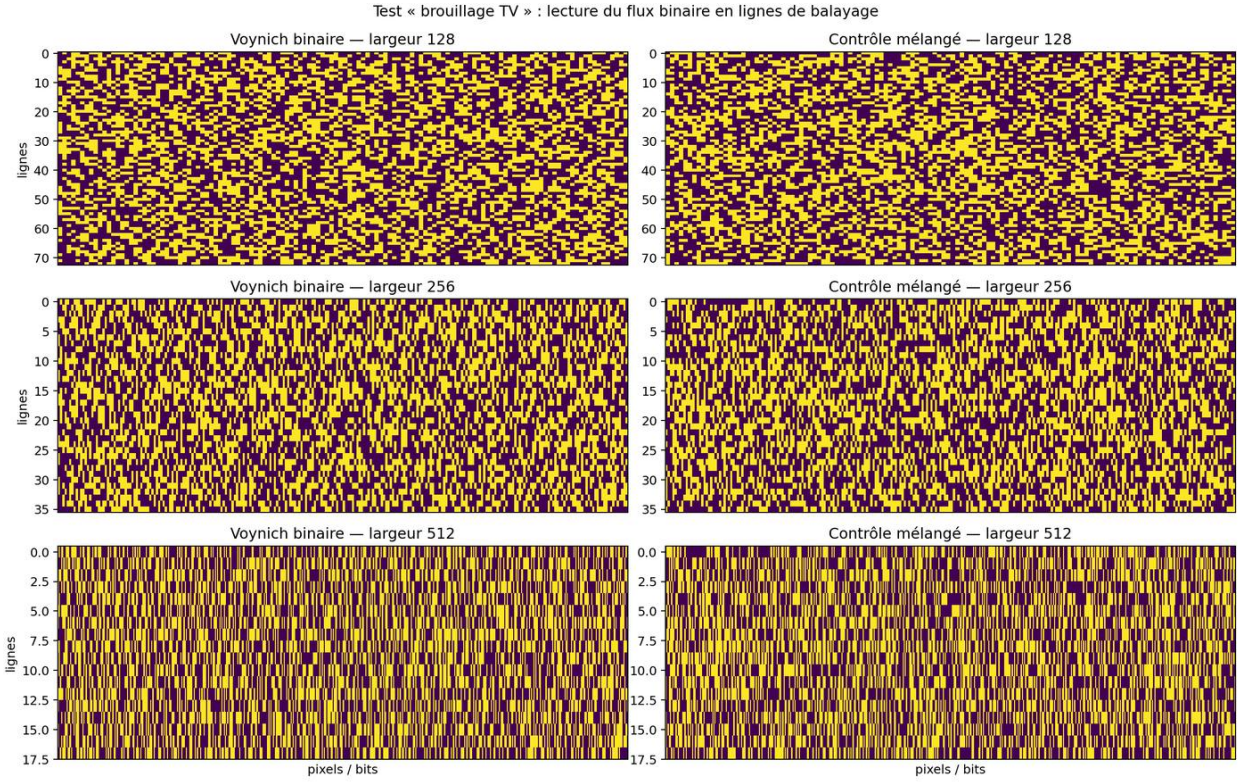

Figure 33 — Video-scrambling-type representation experiment.

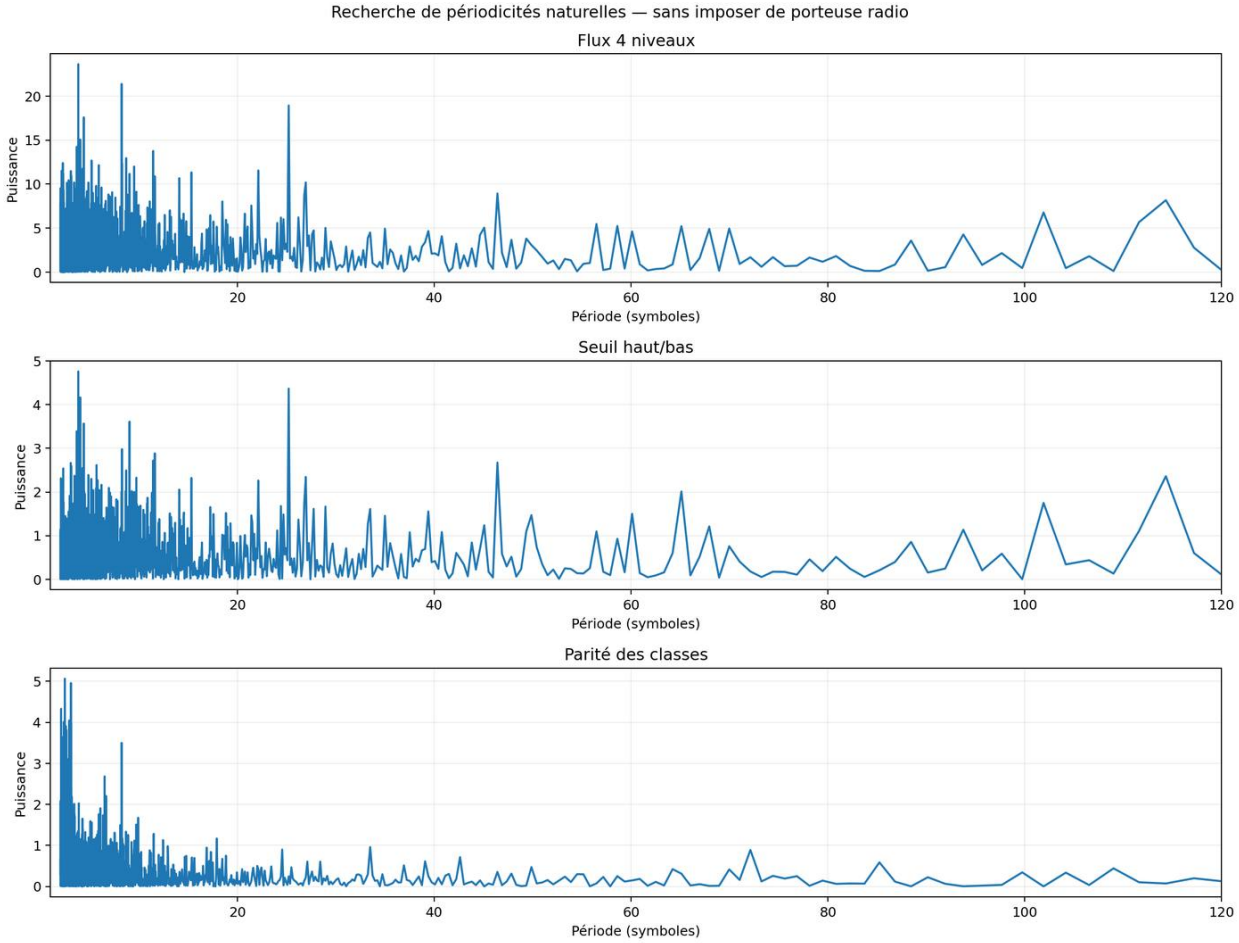

Figure 34 — Exploratory search for a natural carrier.

The idea was to test whether information might be hidden by incorrect synchronization or an incorrect line width, like a scrambled image or signal. No stable content was recovered.

The experiment nevertheless highlighted an essential principle: strongly changing the representation of data can produce highly convincing patterns even when no message is actually present.

15. Reflection on the Wow! Signal

The comparison with the Wow! Signal is conceptual. The goal is not to claim a historical relationship, but to test a general idea: information can be encoded in a representation that resembles noise or a signal until the appropriate decoding method is applied.

No data from this research demonstrates that the manuscript contains the Wow! Signal or an extraterrestrial transmission. The comparison serves as a framework for thinking about coding, carriers, synchronization and weak-signal detection.

16. Hypothesis of a Non-Human or Extraterrestrial Origin

The extraterrestrial hypothesis formed part of the initial questioning and partly explains the search for unconventional codes, signals, astronomical structures and multi-modal representations.

It cannot, however, be inferred from the current results. Even a future discovery of an unusual code would not, by itself, identify its origin. Independent evidence concerning the support, dating, technology, historical context or provenance would be required.

> **Cautionary rule: an unexplained code is not synonymous with an extraterrestrial origin.**

17. Results Summary

Line of inquiry	Observed result	Current status
Binary	Visual regularities, bands and densities	Exploratory; no demonstrated message
Periodicity	Local repetitions and structures	No unique signature
FSK	Constructible and analyzable audio signal	Technical demonstration, not historical signal
ASCII	Partial printable characters	No coherent message
Coordinates	Irregular clouds/trajectories	No clear map
DNA	A/C/G/T sequences and possible ORFs	Hypothesis; random comparison required
Golden ratio	Inconclusive measured ratios	Not demonstrated
Spiral	Non-golden circular fit	Not demonstrated
Astronomy	Real celestial structures; Pleiades plausible	Serious avenue for further study
Light Language	Graphical similarities	No linguistic correspondence
Wow! Signal	Conceptual analogy	No demonstrated link
Extraterrestrial origin	General hypothesis	No evidence

18. What the Research Actually Establishes

- Digital transformations of the Voynich can reveal structures and regularities.
- f115/f116 data can be transformed into binary and then into an audio signal.
- A two-frequency FSK experiment can be constructed from a binary sequence.
- The ASCII attempts did not produce clear, reproducible text.
- A four-state conversion mechanically produces DNA-like sequences.
- ORF-like motifs can appear, but must be compared with random sequences.
- The geometric tests did not confirm the golden ratio.
- Astronomical representations form a distinct and particularly interesting branch.
- Graphic similarities with modern systems are not sufficient to establish historical lineage.

19. What Has Not Been Demonstrated

- The manuscript has not been demonstrated to be extraterrestrial.
- No coherent ASCII message has been recovered.
- No historical radio signal has been extracted from the manuscript.
- The DNA sequence has not been demonstrated to be real biological information.
- No link to the Wow! Signal has been established.
- The golden ratio has not been confirmed by the measurements performed.
- A complete celestial map has not been decoded.
- “Light Language” has not been identified as the language of the manuscript.

20. The Decisive Scientific Next Step

The next phase should move from pattern hunting to measuring improbability. The objective is no longer merely to find an image, sequence or correspondence, but to determine whether the result is statistically exceptional.

20.1. Recommended protocol

• Use a standardized transcription of the manuscript and preserve transcription uncertainties.

• Define conversion rules before inspecting biological or ASCII outcomes.

• Systematically test all 24 DNA mappings.

• Test all predefined FSK offsets and variants.

• Construct random control corpora preserving character and word-length distributions.

• Compare entropy, autocorrelation, periodicity, ORFs, motif frequencies and symmetries.

• Correct results for the total number of tests performed.

• Reproduce any anomalies on independent folios.

• Publish negative results as well, to reduce selection bias.

20.2. Statistical hypothesis

The most useful formulation becomes: “Are the observed properties of the Voynich compatible with a text or symbol sequence having the same distributions, or do they present a statistically reproducible anomaly?”

This formulation allows the linguistic, cryptographic, digital, biological and astronomical hypotheses to be compared objectively without selecting one in advance.

21. Reproducibility and Production Inventory

The experiments generated a set of artifacts documenting the different stages. Among them:

• voynich_115_116_text_binary.png — binary representation of the text.

• voynich_115_116_binary_rewrap.png — binary reshaping.

• voynich_115_116_binary_density.png — binary density.

• voynich_115_116_binary_experiment_montage.png — binary experiment montage.

• voynich_209_pages_periodicites_recurrentes.png — periodicity analysis.

• voynich_50_pages_geometrie_scan.png — geometric scan.

• voynich_f115_116_spectrogramme_audio.png — audio spectrogram.

• voynich_FSK_spectrogramme.png — FSK spectrogram.

• voynich_FSK_decode_binaire.png — binary decoding.

• voynich_ascii_scan.png — ASCII search.

• voynich_ascii_coordonnees_8_offsets.png — ASCII coordinates.

• voynich_115_116_adn_24_mappings.png — 24 DNA mappings.

• voynich_115_116_adn_statistique.png — DNA statistics.

• rosette_nombre_or_mesures.png — golden-ratio measurements.

• rosette_test_spirale.png — spiral test.

• f68r3_big.png / f68r3crop.png / f68r3_detected.png — astronomical analysis.

• pleiades_fit.png — exploratory Pleiades fit.

The audio files include a mono 44.1 kHz/16-bit PCM WAV version and several MP3/WAV variants used to check the robustness of the audio chain.

22. Research Chronology

Phase	Work
1	Questioning the nature of the writing system and the possibility of unconventional coding.
2	Selection of f115/f116 as the experimental area.
3	Binary, raster, density and reshaping conversions.
4	Search for periodicities and recurrent structures.
5	Audio conversion and FSK experiment.
6	Synchronization, offsets, ASCII, UTF-8, hexadecimal and Base64.
7	Four-state transformation and DNA hypothesis.
8	ORF search and comparison with random sequences.
9	Geometric research, symmetries, circles, spirals and golden ratio.
10	Astronomical analysis of f68r3 and the Pleiades hypothesis.
11	Cautious comparison with Light Language and reflection on the Wow! Signal.
12	Construction of a statistical validation protocol.

23. General Conclusion

The research has not yet decoded the Voynich Manuscript. It has, however, established an original experimental approach that treats the document as an information source capable of being represented in many different ways.

The binary, FSK, ASCII and DNA branches show that it is technically easy to produce alternative representations rich in patterns. This is both an opportunity and a danger: without statistical control, a coincidence can appear significant.

The astronomical branch deserves particular attention because it starts from elements explicitly drawn in the manuscript and allows external tests against real celestial positions. The DNA branch can also be retained as a coding experiment, provided it undergoes strict statistical validation.

The next objective should therefore be less spectacular but much more decisive: establish whether the observed anomalies are genuinely improbable under properly constructed null models.

> **In summary: the research has produced several signals of interest, but no decoding proof. The next step is reproducible statistical validation.**

RESEARCH DOSSIER

The Voynich Manuscript — Complete experimental research synthesis

Eric GARCIA

Expert in Digital Research and AI

English edition — September 2026

Translated and reformatted from the complete French research synthesis. The document preserves the exploratory status of the hypotheses and distinguishes them from established results.