STRUCTURAL GRAMMAR OF THE VOYNICH MANUSCRIPT: CHARACTERIZATION AND MECHANISM DISCRIMINATION

A PREPRINT

Asa Shepard
Williams College
as66@williams.edu

April 19, 2026

ABSTRACT

The Voynich Manuscript (Beinecke MS 408), radiocarbon-dated to 1404–1438 CE, contains roughly 38,000 tokens of an undeciphered script that has resisted analysis for over a century. We report a statistical investigation comprising structural characterization and generative-mechanism testing. Characterization establishes five distinct levels of structured grammar in Voynichese, including a three-layer positional word-class architecture, a dialect-inversion finding so sharp that a single word class is the most common line-opener in one dialect and the most common line-closer in the other, and a directed paragraph-level succession chain that cannot be reproduced by chance in either of two independent null models. Mechanism testing across $\sim$5,400 simulations of published and constructed cipher models finds that no position-blind or substitution-only mechanism reproduces the joint structural profile; a position-aware copy-buffer hybrid is the surviving viable class. We confirm this class is non-empty by constructing one such hybrid and verifying it passes a pre-registered specificity test against Latin and Italian corpora. All five structural findings replicate under three independent transliteration systems. We do not claim decipherment or identify a source language; we narrow the space of viable generative hypotheses and characterize the constraints the surviving class must satisfy.

1 Introduction

The Voynich Manuscript, radiocarbon-dated to 1404–1438 CE (Hodgins et al., 2011), contains roughly 38,000 tokens of an unknown script (Voynichese) that has resisted decipherment for over a century. Three broad generative hypotheses dominate the computational literature: that Voynichese encodes natural language through a cipher (with the Naibbe homophonic cipher of Greshko (2025) as the most recent specific proposal); that it is meaningless text generated by a mechanical procedure (with the self-citation algorithm of Timm and Schinner (2020) as the most developed); or that it is a structured system—constructed language, mnemonic code, or otherwise—whose regularities reflect intrinsic rules.

This paper does not adjudicate among these hypotheses. It narrows the space of viable mechanisms and contributes what is, to our knowledge, the most detailed positional and sequential characterization of Voynichese available. The work proceeds in three stages. *Characterization* (Section 2) establishes that Voynichese exhibits structured grammar at five distinct levels, with the terminal-layer dialect inversion and the directed A-herbal paragraph chain as the sharpest novel findings. *Discrimination* (Sections 3–4) establishes a dual constraint that no position-blind cipher, no substitution-only cipher, and no naïve copy-buffer hybrid can satisfy. *Constructive existence* (Section 5) reports a position-aware copy-buffer hybrid that does satisfy the joint profile, establishing that the surviving mechanism class is non-empty and producing one falsifiable mechanistic finding (C009) about copy-buffer limits. The construction is not claimed as evidence that this mechanism class is the right answer for Voynichese; it is a feasibility demonstration with explicit degrees-of-freedom disclosure.

We make a methodology disclosure prominently. The work was conducted by an autonomous multi-agent loop with a Sonnet-class agent running experiments and an Opus-class agent performing independent meta-review at phase boundaries; all strategic decisions and writing are the author's. Five substantive corrections produced by the meta-review process are relevant to the claims here, including a target-awareness disclosure on a position-sensitive cipher, an identification that one hybrid's modification operation was maximally unfavorable to the hypothesis being tested, and a misattribution catch in committed-claim text. Each correction was in the direction of overstatement relative to evidence; each was caught externally rather than by the primary agent's self-checking. We treat this as encouraging but not as the contribution of this paper.

We do *not* claim decipherment, identify a source language, or rule out untested mechanism classes (transposition ciphers, running-key ciphers, position-sensitive ciphers of non-Latin source languages, structured glossolalia). We claim that the tested mechanism classes are incompatible with Voynichese's joint structural profile in the specific and quantified ways reported, and that the surviving mechanism class is non-empty.

2 Voynichese Structural Grammar

Existing characterizations (Currier, 1976; D'Imperio, 1978; Stolfi, 2000; Zattera, 2022; Bowern and Lindemann, 2021) establish that Voynichese has a Currier A/B dialect split—two statistically distinct populations of folios first identified by Currier 1976—a productive suffix paradigm, a low-high-low entropy profile within words, and slot-grammar structure. Building on ∼579 corpus-statistical tests against the EVA ZL3b transliteration (Zandbergen, 2023)—EVA (Extended Voynich Alphabet) being the standard scholarly system for rendering the manuscript's script in Roman letters—we extend these in five directions.

A note on terminology used throughout this section. Voynichese words are grouped by *onset class*—the leading letter sequence in EVA notation (for example, words beginning *d*-, *sh*-, or *qok*-)—into roughly 20 functionally distinct categories that act as word-class proxies. *Positional enrichment* is a ratio: 1.0 means a class appears at a given line position exactly as often as chance predicts; 5.23× means more than five times as often. A result is *confirmed* (CONF) if it clears a conservative statistical threshold correcting for the large number of comparisons made (Bonferroni correction at m=175 tests, minimum 10 observations). A *null model* is a randomly shuffled version of the data used to establish what patterns could arise by chance alone. All five findings below were *pre-registered*—their methods and acceptance thresholds committed to a timestamped record before experiments ran—to guard against post-hoc adjustment.

Three-layer positional architecture (Figure 1). Onset classes partition by line-position enrichment into openers (confirmed as line-first enriched), mid-classes (neutral at boundaries), and terminals (confirmed as line-final enriched). The B-dialect openers are $\{p\,(5.23\times), s\,(4.77\times), y\,(3.54\times), t\,(3.18\times), d\,(2.64\times)\}$—meaning, for instance, that p-onset words appear at line-first position more than five times as often as chance would predict; mid-classes $\{ch, qok, ok, ot, qot, sh, k\}$; terminals $\{r\,(2.80\times), l\,(1.85\times), ol\,(1.65\times), o\,(1.58\times), a\,(1.52\times)\}$. The A-dialect partition is structurally analogous but with critical inversion: A terminals are $\{d\,(1.53\times), cth\,(1.67\times)\}$, sharing zero classes with B terminals. This terminal-layer inversion is the cleanest dialect discriminant in the manuscript: d-onset is the dominant A line-closer (18.4% of A lines) and the dominant B line-opener (19.2% of B lines). A complete glossary of shorthand used throughout (ss_y, ss_d, r_{lf}, r_{mid}, ZeroMid, β_{scale}, am_{ll}, and related terms) appears in supplementary material.

Three-register line grammar. Lines partition by folio position into openers (first line of each folio, 67–80% gallows-onset LF, mean 7–8 words, 85% hapax line-final vocabulary), continuations (baseline distributions), and closers (last line, anti-gallows LF 0.44–0.60×, mean 4–5 words). Gallows glyphs are the tall ascender characters in EVA (*p*, *t*, *k*, *f*); hapax words are those appearing only once in the corpus. The pattern is section-invariant. The A-dialect p-onset shows complete folio-position exclusion: 37.3% of folio-opener line-final tokens, 0 of 83 folio-closer line-final tokens. The gallows-opener formula is a one-word constraint operating on word-index 0 only, with no internal template structure—the line-final vocabulary is dominated by words appearing only once.

Bigram grammar with cross-dialect invariants (Figure 2). A bigram is a pair of successive words; here we track which onset class tends to follow which other. The complete 7×7 onset-class succession matrix, computed relative to a chance baseline, yields five cross-dialect invariant elevated pairs—pairs that follow each other more often than expected in both dialects ($y{\to}y$ at 2.49×/2.85×, $sh{\to}qok$ at 1.97×/1.74×, $ok{\to}ok$ at 1.32×/1.96×, $qok{\to}s$ at 1.36×/1.39×)—and one cross-dialect avoidance, meaning the two classes follow each other far less often than chance ($sh{\to}ch$ at 0.555×/0.565×). B-dialect adds bidirectional $ch{\leftrightarrow}sh$ avoidance ($ch{\to}sh = 0.567\times$). A-dialect adds $qok{\leftrightarrow}ok$ bidirectional coupling (2.03×/1.68×) absent in B. The $sh{\to}qok$ coupling is driven by the she- body subclass (*shedy*, *shey*, *sheedy*) operating as a triple-hub node receiving $d{\to}sh$ and $s{\to}sh$ and emitting $sh{\to}qok$.

Paragraph-level chain grammar (Figure 3). A herbal sections exhibit a directed acyclic cross-line succession network: $d \to qot \to ch \to o \to sh$, with t as a relay class receiving from d and sh. The chain operates strictly at

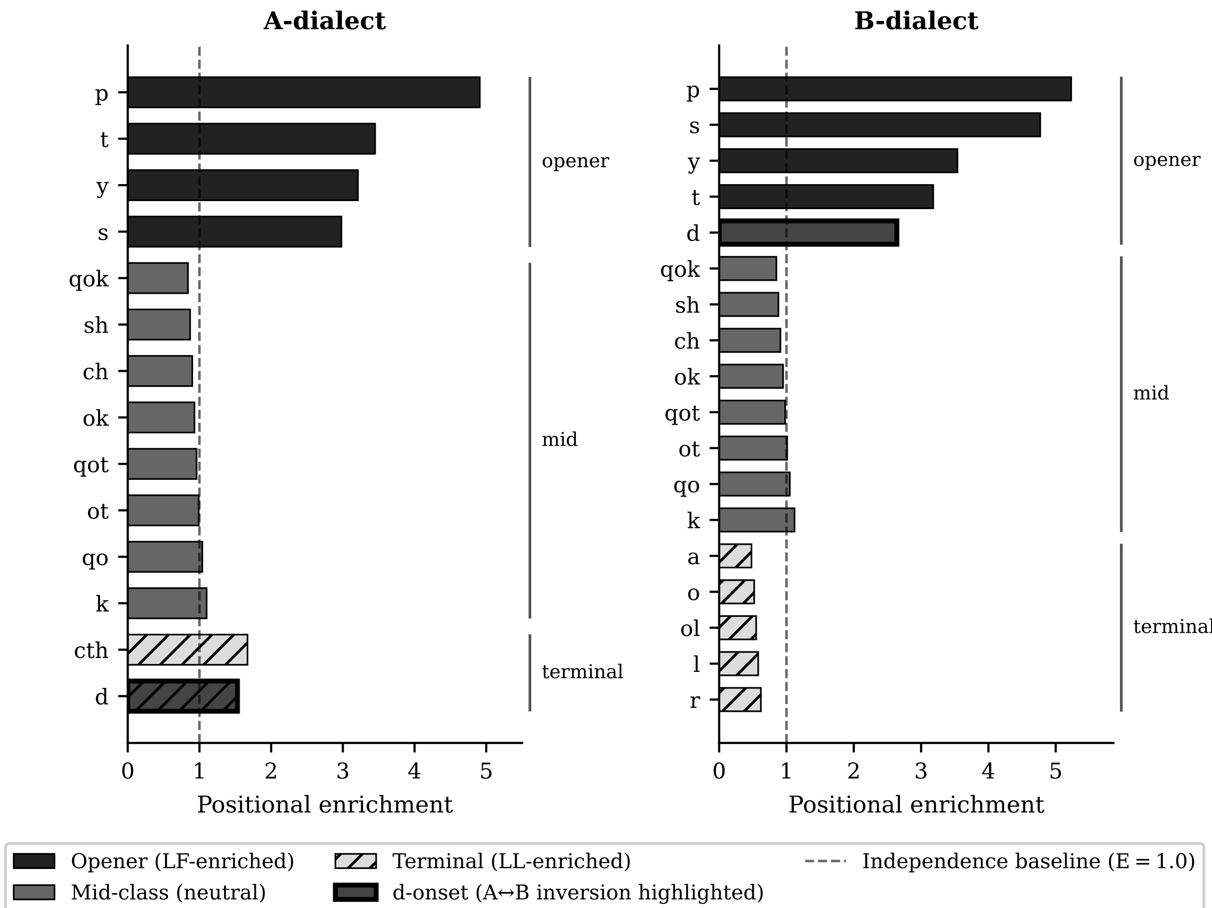

Figure 1: Three-layer positional architecture. Horizontal bars show positional enrichment for each onset class in A-dialect (left) and B-dialect (right). Openers (dark) are enriched at line-first position; terminals (hatched) are enriched at line-final position; mid-classes are neutral at both boundaries. The **d**-onset (highlighted) is the dominant B-dialect opener and the dominant A-dialect terminal—the sharpest cross-dialect inversion in the data.

the inter-line level: within-line $d{\rightarrow}qot$ succession is null ($A : 0.99\times$, $B : 0.66\times$ depleted), while inter-line $d{\rightarrow}qot$ is CONF. This is the clearest scope-separation signal in the data—two distinct mechanisms operate at different structural levels. The chain is A-herbal-specific; B-dialect cross-line patterns are heterogeneous and section-specific (B astro: qok self-chains; B pharma: $ol{\rightarrow}qok$ at $2.76\times$). Under null-model testing—that is, comparing the chain against 200 versions of the data with sections randomly shuffled and 200 versions with onset labels randomly shuffled—0 of 200 controls in either case reproduce the chain's joint per-edge strength. The chain lies at the 100th percentile of both null distributions. We report the chain as a theoretically-motivated directed finding specified a priori by the bigram characterization in the previous paragraph, not as a globally-strongest chain identified by exhaustive search; this distinction matters because exhaustive-search statistics over all directed length-5 paths are dominated by sparse-class pathology, while the specified chain survives the proper null.

Systematic dialect transformation. Across every onset class, A-dialect uses $-ol$/$-or$/$-ey$ suffixes and B-dialect uses $-edy$/$-eedy$. This is the most systematic vocabulary transformation between dialects, affecting 20+ characterized onset classes uniformly. Combined with the terminal-layer inversion and the structural identity $\text{adj\_rate}_A = \text{adj\_rate}_B = 0.00933$ (matching to four decimal places), this points to systematic encoding difference rather than scribal drift.

Transliteration replication. All five characterization findings were replicated under three alternative transliterations (EVA-basic, v101 Claston, and ZL3a) against pre-registered thresholds committed to the project repository before the experiments ran. Table 1 reports the per-finding, per-transliteration results. Four of five findings replicated in all three systems; the A-herbal paragraph chain (Finding 4) replicated under EVA-basic (96th/97.5th percentile) and ZL3a

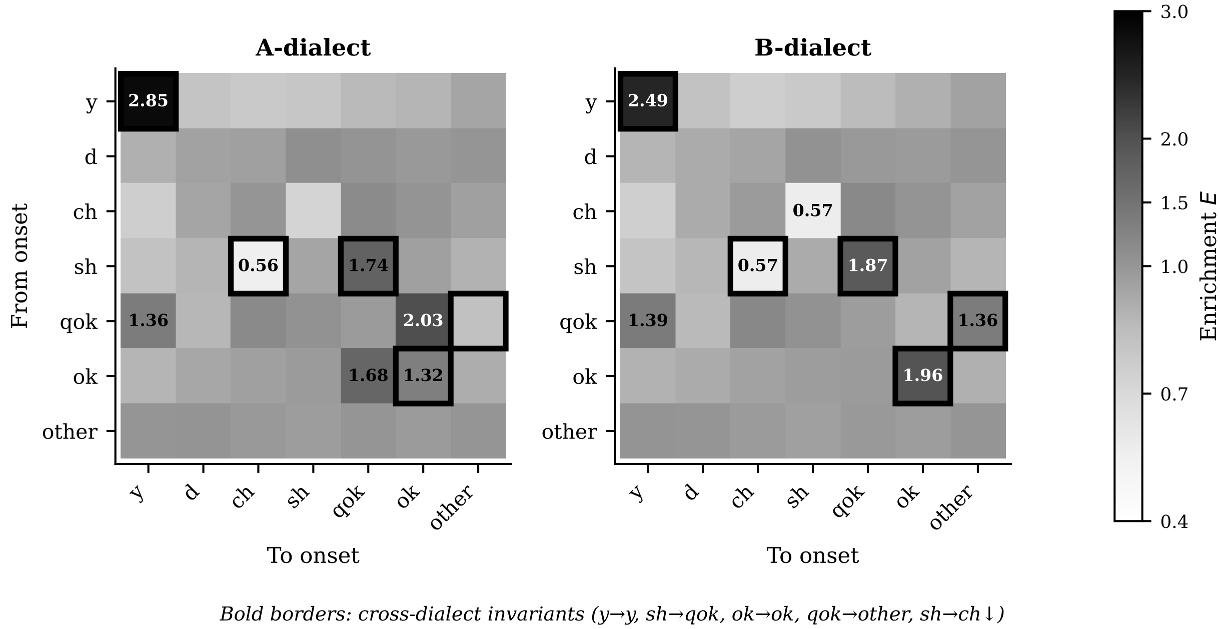

Figure 2: Onset-class bigram succession matrix for A-dialect (left) and B-dialect (right). Shading encodes enrichment E (dark = elevated, white = depleted, mid-gray ≈ 1); values are annotated where $E \geq 1.3$ or $E \leq 0.7$. Bold borders mark the five cross-dialect invariants: $y{\to}y$, $sh{\to}qok$, $ok{\to}ok$, $qok{\to}$other (B), and $sh{\to}ch$ depletion.

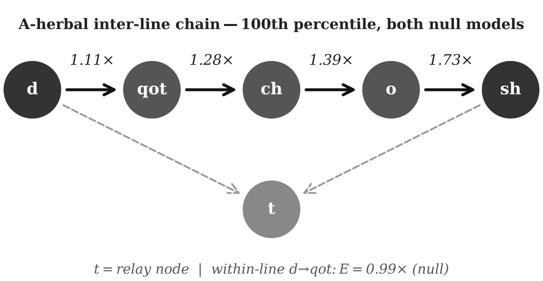

Figure 3: A-herbal paragraph-level chain ($d \to qot \to ch \to o \to sh$, with t as relay). Edge labels show inter-line enrichment values. Note that within-line $d{\to}qot$ is null ($E = 0.99\times$); the chain operates strictly at the cross-line level. The chain lies at the 100th percentile of both section-shuffled and onset-shuffled null distributions ($0/200$ controls in each case).

(97.5th/96.5th percentile) but fell below the pre-registered 95th-percentile threshold under v101 (77th/86.5th percentile), attributable to reduced observation counts on the $o{\to}sh$ chain edge under v101's tokenization (obs $= 1$ vs obs $= 4$ under ZL3b). Per the pre-registered rule requiring retention in ≥ 2 of 3 alternative transliterations, all five findings are classified as retained. The v101 partial failure on Finding 4 indicates sensitivity of the fine-grained paragraph chain to transliteration-specific onset-class boundaries and is noted as a limitation.

A separate finding bears on interpretive claims about specific Voynich words. The token *daiin* accounts for 27.4% of all d-onset tokens. Its dominance is fully explained by the convergence of d-onset frequency, the d-onset $\to -aiin$ body constraint (34.6% A, 22.5% B), and its position-neutral status (ZeroMid: enriched at neither line boundary)—it requires no special semantic explanation. Interpretations treating *daiin* as a key content word should account for this mechanistic baseline.

Table 1: Transliteration replication of Section 2 findings. Retained/Failed per pre-registered thresholds (committed before experiments). Finding 4 under v101 falls below the 95th-percentile null-model threshold due to reduced observation counts on one chain edge.

Finding	EVA-basic	v101	ZL3a	Overall
F1: Three-layer architecture	18/18 R	16/18 R	18/18 R	**Retained**
F2: Terminal inversion	R	R	R	**Retained**
F3: Bigram invariants	4/5 R	4/5 R	4/5 R	**Retained**
F4: Paragraph chain	R	**F**	R	**Retained (2/3)**
F5: Suffix transformation	20/21 R	19/22 R	20/21 R	**Retained**

3 Mechanisms Tested

We evaluate generative mechanisms against the Currier B corpus (2,447 lines, 20,490 tokens) under matched line-length distributions. Reported confidence intervals are empirical 2.5–97.5 percentiles.

Three broad hypotheses have been proposed for Voynichese: that it encodes natural language through a cipher, that it is procedurally generated text with no underlying language, and that it is some form of structured hybrid. We test one or more specific mechanisms from each class by running simulations and measuring how well their output matches the statistical profile established in Section 2.

Naibbe (Greshko, 2025): a homophonic substitution cipher in which each plaintext character can map to multiple possible ciphertext characters, recently proposed as a specific generative model for Voynichese. Tested at published parameters, applied to Italian (Dante) and Latin (Vulgate) plaintexts under two line-breaking protocols (400 runs).

Timm self-citation (Timm and Schinner, 2020): an algorithm that generates new text by copying and lightly modifying words from nearby positions in the growing output—proposed as a mechanism that can produce Voynich-like word distributions without encoding any natural language. Tested across a 75-point parameter sweep (COPY_PROB $\in \{0.3, 0.5, 0.7, 0.85, 0.95\}$, MODIFY_PROB $\in \{0.1, 0.25, 0.4, 0.6, 0.8\}$, WINDOW $\in \{10, 30, 50\}$; 1,500 runs).

Position-sensitive substitution cipher: a cipher that applies different character substitution tables depending on whether a word falls at the start, middle, or end of a line, designed to test whether position-dependent rules can reproduce Voynichese's positional structure. Applied to Cicero *De Officiis*. We disclose that the line-final table is target-aware: all Latin clause-final morphemes ($-am$, $-um$, $-us$, $-em$, $-is$, $-ae$, $-orum$, $-bus$, $-unt$, $-it$, $-ant$, $-at$) map to EVA $-am$ specifically, because $-am$ is the metric target. This makes the resulting null finding stronger, not weaker.

Hybrid architectures. Four configurations: position-blind copy buffer over the cipher (exp_179[1]); position-aware buffer restricted to mid-line with onset-replacing modification (exp_180); position-aware buffer with onset-preserving suffix-replacement modification (exp_181); position-aware buffer with Timm-style character-level modification preserving onset (exp_182). Total: 4,400+ hybrid simulations.

A position-aware copy-buffer hybrid with full calibration. A 21-mechanism variant in which line-first and line-final words are generated by a position-sensitive onset/suffix mechanism, mid-line words are drawn from a rolling copy buffer with onset-preserving character-level modification, and additional mechanisms perform vocabulary clamping; Markov onset coupling (in which each onset class's probability depends on the preceding class, reproducing the bigram structure from Section 2); n-conditional Menzerath scaling (an empirical linguistic pattern in which longer words tend to occur alongside shorter ones, applied here to match Voynich word-length distributions); and short-line line-final word-length filtering. Six parameters are explicitly target-calibrated. Full enumeration of mechanisms and parameter values appears in supplementary material. This variant is treated as a constructive existence proof (Section 5) rather than as a further mechanism under discrimination test.

4 Discrimination: The Dual Constraint

Two metrics drive the discrimination. *Self-succession* measures how often a given onset class follows itself on adjacent words relative to chance: 1.0 means as often as chance predicts; $1.47\times$ means 47% more self-following than expected. The *adjacent identical word rate* (adj_rate) is the fraction of consecutive word pairs that are exact repeats. $-am$ *line-final enrichment* is the positional enrichment of words ending in the $-am$ suffix at the last position of a line—Voynich B's strongest positional signal.

[1]Experiment identifiers (exp_*NNN*) reference per-experiment results in supplementary material.

Table 2: Mechanism failure matrix. **Bold** indicates the value falls outside the Voynich-consistent range. PS-cipher uses the target-calibrated strong setting.

Metric	Voynich B	Naibbe	Timm (sweep)	PS-cipher
$-am$ line-final enr.	6.811	**1.008**	**0.572**	**5.756**
y-LF	3.506	**1.011**	**0.526**	2.395
Self-succession (mean)	1.465	**0.92**	1.5–2.4	**0.999**
Adj. rate	0.009	≈ 0.009	**0.037**	≈ 0.009

Position-blind mechanisms fail on positional grammar. Naibbe produces $-am$ line-final enrichment of $1.008\times$ (vs. Voynich $6.811\times$); the Timm sweep maximum across 75 combinations is $0.572\times$. The convergent failure of two architecturally distinct mechanisms at the same metric establishes that we are ruling out the *class* of position-blind mechanisms, not specific implementations.

Substitution alone produces no sequential memory. The target-calibrated position-sensitive cipher reaches $5.756\times$ on $-am$ line-final enrichment (84% of Voynich) but produces self-succession of exactly $0.999\times$—the independence expectation. This is not a calibration issue; substitution ciphers operate token-wise.

Timm produces sequential memory at wrong magnitudes. Across the parameter sweep, no setting jointly produces Voynich-level self-succession ($\sim 1.47\times$) and Voynich-level identical-adjacent rate (0.009). At published parameters, Timm produces self-succession 1.5–$2.4\times$ but adj_rate 0.037 (fourfold excess). Reducing COPY_PROB lowers both metrics in lockstep.

Hybrid progression. The position-blind hybrid (exp_179) fails because copying at line-final positions dilutes the positional signal linearly: each $+0.1$ in COPY_PROB costs -0.47 in $-am$ line-final enrichment. Position-aware gating (exp_180) decouples the positional signal ($-am$ line-final enrichment flat at 5.7–$5.8\times$ across all COPY_PROB) but onset-replacing modification destroys the self-succession feature, producing $0/18$ jointly viable combinations. Suffix-replacement modification (exp_181) decouples all three metrics and reaches $4/36$ jointly viable, with best-combination adj_rate 0.0144 ($1.6\times$ above Voynich). Character-level modification preserving onset (exp_182) reaches $3/36$ viable at the strict adj_rate ≤ 0.009 threshold—the first mechanism to cross the Voynich exact-repetition rate while maintaining the positional and sequential signals.

Table 2 and Figure 4 summarize. The dual constraint defines minimum requirements for any viable generative hypothesis; one mechanism class (position-aware copy-buffer hybrid with morphologically plausible modification) is shown by exp_182 to contain at least one satisfying configuration.

5 Constructive Existence: A Position-Aware Copy-Buffer Hybrid Reaches the Joint Profile

The discrimination result in Section 4 narrows viable mechanisms, among the mechanism classes tested, to a position-aware copy-buffer hybrid class with onset-preserving modification; other mechanism classes not evaluated here may also satisfy the joint constraints. To establish that this class is non-empty at the full joint profile, we construct one such hybrid and verify it reaches all 19 metrics of a B-dialect battery within $\pm 20\%$ tolerance under target calibration ($N{=}30$). Structurally, the hybrid generates line-first and line-final words by position-sensitive onset/suffix mechanisms and mid-line words by a rolling copy buffer with onset-preserving character-level modification; supplementary mechanisms perform vocabulary clamping; Markov onset coupling ($p_{qok \to ch} = p_{ch \to qok} = 0.12$), in which each onset class's probability depends on the preceding class (see Section 3); and n-conditional Menzerath scaling (see Section 3). Six parameters are target-calibrated; the remainder are structural. The construction demonstrates that the dual constraint is satisfiable—that some member of the surviving class attains the joint profile—but is not intended as evidence that this mechanism class is the generative model of Voynichese.

Degrees of freedom should be disclosed explicitly. The architecture has 21 component mechanisms with 6 target-calibrated parameters against a 19-metric battery whose effective independent dimension is lower than 19. Under PCA on the cross-simulation metric covariance, 16 components capture 95% of variance; under hierarchical correlation clustering at $|r| \geq 0.7$, the battery resolves to 18 independent clusters. One pair is mechanically linked at $r = -0.94$ (am_{ll} and $am_{\mathrm{mid\_enr}}$, jointly determined by the line-final suffix mechanism). The full 19×19 metric-correlation matrix is in supplementary material. Even at the more favorable reading (18 effective constraints against 6 calibrated parameters), the architecture is not overdetermined, and we do not claim its $19/19$ battery success is itself evidence for the mechanism class. Under a pre-registered specificity test (exp_585)—thresholds fixed before results were observed, as described in

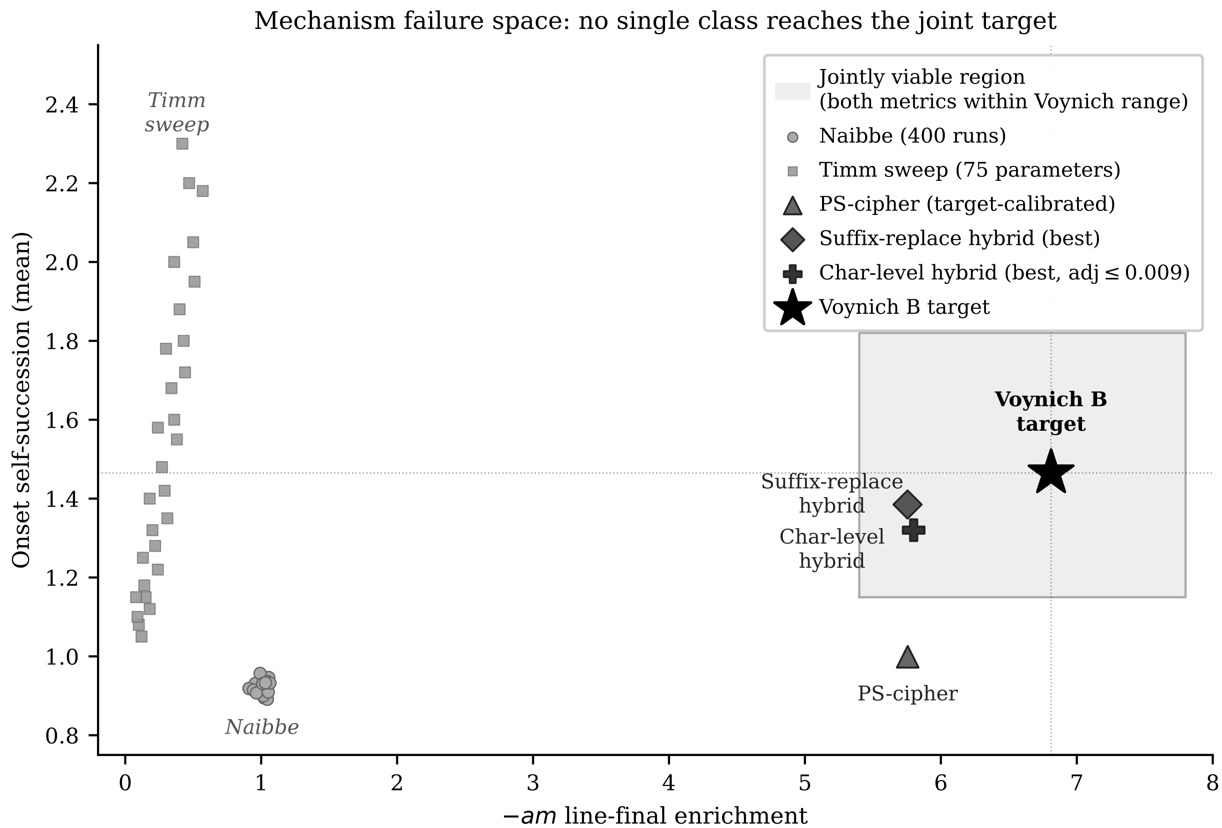

Figure 4: Mechanism failure space. Each tested mechanism class is plotted by $-am$ line-final enrichment (x-axis) and onset self-succession mean (y-axis). The Voynich B target (star) requires satisfying both axes simultaneously. The gray shaded region marks the jointly-viable zone. No position-blind cipher (Naibbe, Timm) approaches the x-axis target; substitution alone (PS-cipher) fails the y-axis; hybrid progression moves monotonically toward the target as architectural constraints are added.

Section 2—the architecture was re-optimized against two structurally-matched non-Voynich corpora (Latin Cicero *De Officiis* and Italian Dante *Divina Commedia*, both segmented to Voynich-matched line-length distributions) and against inter-shuffled Voynich. The architecture achieves $17/17$ applicable metrics on real Voynich, $5/17$ on Latin, $4/17$ on Italian, and $12/17$ on inter-shuffled Voynich, yielding a pre-registered specificity ratio $R = 17/12 = 1.417$ (threshold $R \geq 1.3$ for strong specificity; met). The six structural failures on Latin and Italian are mechanistically grounded: VMS line-first onset distributions, vocabulary-level $-am$ suffix rates, and copy-buffer self-succession floors are built into the architecture's fixed mechanism constants and cannot be calibrated away. This establishes that the architecture's success on Voynich is specific to Voynichese structure rather than a property of the fitting procedure.

The construction does yield one falsifiable mechanistic finding (C009). The architecture overproduces adjacent identical word pairs by 1.4–$1.8\times$ for *sh*/*ch*/*d* classes independent of vocabulary clamping. Diagnosis: Voynich adj_rate for these classes is consistent with random draws from finite vocabulary pools (sh: $1.11\times$ random expectation; d: $0.70\times$; ch: $1.91\times$ via self-succession concentration), and Voynichese has no special anti-repetition mechanism for them; the overproduction arises from copy-buffer/self-succession coupling, because sh (9.2%), ch (15.0%), and d (8.6%) self-succession rates cause any copy-buffer operating on same-onset successors to systematically over-repeat words. The resulting claim—any copy-buffer with $\geq 8\%$ same-onset self-succession overproduces adj_rate for that class, independent of vocabulary calibration—further constrains the surviving mechanism class beyond what the discrimination alone establishes. Separately, s-onset and t-onset show adj_rate $= 0$ in Voynich B ($0/452$, $0/466$ tokens), a structural zero from their positional constraint as line-openers.

A section-aggregation caveat also applies. The B-dialect corpus pools four sections with systematically different positional grammar profiles: bio (am_{ll}=0.053), herbal, astro, pharma (am_{ll}=0.192). Document-order metric drift (first-half $am_{\text{ll}} = 5.825$, second-half 7.147) is 97–107% explained by section composition. The $19/19$ is therefore a

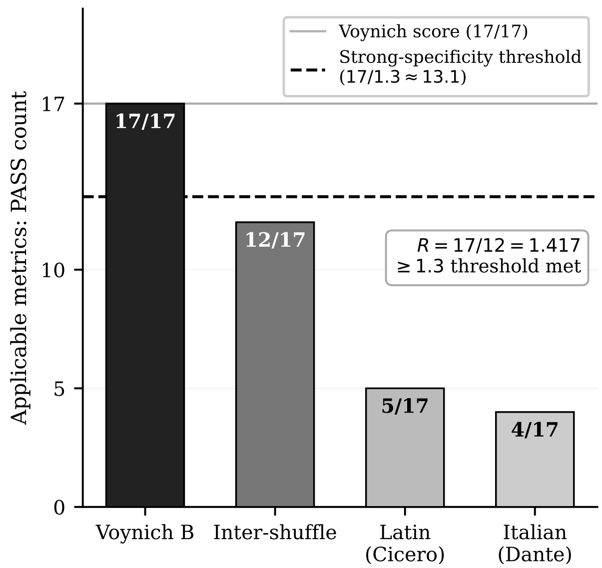

Figure 5: Architecture specificity: applicable battery PASS counts across four test corpora. The dashed line marks the pre-registered threshold below which non-Voynich corpora must fall for strong specificity ($17/1.3 \approx 13.1$). All non-Voynich corpora fall below this threshold; the specificity ratio $R = 17/12 = 1.417$ meets the pre-registered $R \geq 1.3$ criterion.

full-corpus aggregate whose target metrics may not correspond to any single section's underlying statistics. Section-level batteries are the appropriate next test of whether the hybrid captures section-uniform structure or only the aggregate mixture, and have not been performed; until they are, the architecture's success should be read as aggregate-level feasibility only.

6 Methods

We use the EVA ZL3b transliteration (Zandbergen, 2023), filtering to paragraph-type loci, with dialect assignment by per-line $-edy$ rate (canonical methodology; supersedes earlier folio-level assignment). Yielding B-dialect: 40 folios, 2,447 lines, 20,490 tokens. Positional enrichment uses one-sided binomial exact tests, Bonferroni-corrected at m=175. Self-succession is computed as $P(C \mid \text{prev}{=}C)/P(C)$ over in-line adjacency pairs, with stratified position control. Cross-validation is $80/20$ stratified by dialect plus leave-one-section-out. Null-model tests for the A-herbal paragraph chain use 200 section-shuffled and 200 onset-shuffled controls, comparing the joint per-edge enrichment of the specified chain against the null distribution. Code, data, per-experiment directories, and random seeds are available on request; canonical methodology, an executable validator, the 19×19 metric-correlation matrix referenced in Section 5, a glossary of shorthand, and the full enumeration of the 21 architecture mechanisms are included as supplementary material.

Replication thresholds for the transliteration study (Section 2) and specificity thresholds for the architecture test (Section 5) were pre-registered in a timestamped repository commit before the respective experiments ran. The pre-registration document, commit hash, and deviation log are in supplementary material.

The autonomous-loop methodology produced five substantive corrections during this project: (1) retraction of a long-range mutual-information finding contaminated by within-line-shuffle artifacts; (2) target-awareness disclosure on the position-sensitive cipher's line-final substitution table; (3) identification that the exp_180 modification operation was maximally unfavorable to the hypothesis being tested, directly producing exp_181/exp_182; (4) a misattribution catch in committed-claim text (the value $2.05\times$ originally attributed to A ok self-succession actually belonged to B qo); (5) a line-position convention inversion catch during specification audit. Each correction was in the direction of overstatement. We document this for transparency, not as a contribution.

7 Limitations

Untested mechanism classes: transposition ciphers, running-key ciphers, book ciphers, position-sensitive ciphers of non-Latin source languages, and structured glossolalia. The dual constraint rules out tested classes only.

Transcription dependence: All five Section 2 findings were replicated under three alternative transliterations (Table 1). Four of five replicated in all three systems; the paragraph chain weakened under v101 due to transliteration-specific observation-count reduction on one edge. Findings at the suffix-class level remain sensitive to transliteration choice at the margin; the replication establishes robustness of coarse and medium-grained structural findings.

Line-as-unit assumption: positional grammar metrics assume physical lines correspond to linguistic units. This is shared with the literature but unestablished.

Section heterogeneity: all B-dialect aggregate findings are subject to the section-composition caveat in Section 5. Scribal-hand assignments (Davis, 2020) are near-perfectly co-linear with section composition in B-dialect, preventing independent attribution of the gradient to scribal variation.

Architecture degrees of freedom: the constructive existence architecture has 21 component mechanisms and 6 target-calibrated parameters against a 19-metric battery with effective independent dimension of 16–18. Its $19/19$ PASS should be read as constructive existence—a demonstration that the dual-constraint-surviving mechanism class is non-empty—not as evidence that this mechanism class is the correct generative model of Voynichese. A pre-registered specificity test against Latin and Italian corpora returned $R = 1.417$ (Section 5), establishing specificity under the tested non-Voynich substrates. Specificity against other corpus types (e.g., medieval Hebrew, Arabic, or constructed languages with similar morphological complexity) remains untested.

No decipherment, no source language, no unique mechanism identification: even taking the architecture as accurate, multiple mechanism families could potentially satisfy the same constraints.

Methodology under-evaluated: five catches from a single project is encouraging but not a basis for strong claims about the autonomous-loop methodology, which we intend to document separately.

8 Conclusion

Voynichese exhibits structured grammar at five distinct levels: a three-layer positional architecture with complete A/B terminal-layer inversion, a section-invariant three-register line grammar, a 7×7 bigram matrix with cross-dialect invariants, a directed paragraph-level chain in A herbal sections that survives onset-shuffled and section-shuffled null models at the 100th percentile, and a systematic dialect-marking suffix transformation. These findings replicate under three alternative transliterations (EVA-basic, v101 Claston, ZL3a) against pre-registered thresholds, with four of five findings retained in all three systems. These are the contributions on which the paper rests. A dual structural constraint—joint position-sensitive onset-class grammar and sequential word-class memory—rules out position-blind ciphers, substitution-only ciphers, and naïve copy-buffer hybrids in $\sim$5,400 simulations across published and constructed mechanisms. The surviving viable mechanism class must be jointly position-aware, sequentially structured at the onset-class level, and capable of producing onset-class clustering through a mechanism that does not impose excess exact-word repetition; a position-aware copy-buffer hybrid with onset-preserving character-level modification reaches the joint profile, establishing non-emptiness of the surviving class and yielding one mechanistic constraint (C009): any copy-buffer with $\geq 8\%$ same-onset self-succession overproduces adjacent identical words for that class. The architecture's specificity to Voynichese is established by a pre-registered test against Latin and Italian corpora ($R = 1.417$, threshold ≥ 1.3). Whether the surviving class is best instantiated as an intrinsic structured grammar (constructed or natural language), as a copy-buffer hybrid of some form, or as a mechanism outside what we tested, is a question outside what these experiments can settle.

References

Claire Bowern and Luke Lindemann. The linguistics of the Voynich manuscript. *Annual Review of Linguistics*, 7:285–308, 2021.

Prescott H. Currier. Some important new statistical findings. Voynich manuscript conference proceedings, 1976.

Lisa Fagin Davis. How many glyphs and how many scribes? Digital paleography and the Voynich manuscript. *Manuscript Studies*, 5(1), 2020.

Mary E. D'Imperio. *The Voynich Manuscript: An Elegant Enigma*. National Security Agency, 1978.

Michael Greshko. The Naibbe cipher: a substitution cipher that encrypts Latin and Italian as Voynich Manuscript-like ciphertext. *Cryptologia*, 2025.

Jorge Stolfi. A grammar for Voynichese words. Technical report, Universidade Estadual de Campinas, 2000.

Greg Hodgins et al. Radiocarbon dating of the Voynich Manuscript. University of Arizona AMS Laboratory, 2011.

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

René Zandbergen. The EVA transliteration alphabet. `http://www.voynich.nu/`, 2023.

Massimiliano Zattera. A new transliteration alphabet brings new evidence of word structure and multiple languages in the Voynich manuscript. International Conference on Historical Cryptology, 2022.