One Hand, Five Labels: A Critical Examination of the Five-Scribe Hypothesis for the Voynich Manuscript

Torsten Timm

ARTICLE HISTORY
Compiled March 2, 2026

ABSTRACT
The five-scribe hypothesis proposed by Davis (2020c) has become a foundational premise in recent Voynich Manuscript research, influencing linguistic analysis, topic modeling, and public understanding of the manuscript. This paper subjects the hypothesis to critical examination.

Building on the continuous evolution framework established in Timm (2026a), which demonstrated that the manuscript's text evolves gradually from one state to another rather than partitioning into discrete systems, this paper shows that: (1) Davis's diagnostic criteria—the glyphs <k> and <n>—fail on empirical examination, with both variants appearing on nearly every page of the manuscript; (2) the five scribes reduce to pre-existing categories, with Scribes 1 and 2 replicating the Currier A/B language distinction and Scribe 4 identifying pages with labels rather than running text; (3) the chain of apparent independent confirmation is circular, with each study inheriting assumptions from the previous one; and (4) the handwriting variation is more parsimoniously explained as continuous evolution of a single hand alongside the documented continuous evolution of the vocabulary. The conclusion is that discrete scribe categories, like discrete language categories, are imposed on a continuum.

1. Introduction

A companion paper (Timm, 2026a) argued that the persistent failure of cryptographic, linguistic, and statistical approaches to the Voynich Manuscript stems from a shared foundational assumption: the manuscript's text was produced by a static system. The evidence presented there—a single connected word network spanning 84.67% of word types, systematic correlation between word frequency and number of similar word forms, and smooth cosine similarity gradients with no abrupt transitions between sections—demonstrated that the text evolves continuously throughout the manuscript rather than partitioning into discrete systems. This continuous evolution is consistently explained by the self-citation hypothesis (Timm and Schinner, 2020), which proposes that the text was generated through iterative copying and modification of previously written words.

If the vocabulary evolves continuously, a natural question arises: does the handwriting evolve in parallel? The five-scribe hypothesis proposed by Lisa Fagin Davis claims otherwise. On the basis of paleographic analysis, Davis (2020c) identifies five clearly distinguishable scribes in the manuscript and assigns specific folios to each. This hypothesis has had substantial influence on subsequent research. Bowern and Lindemann

(2021) treat the five scribes as established when analyzing the manuscript's linguistic properties. Sterneck et al. (2021) report correlations between computationally derived topic clusters and Davis's scribe assignments, interpreting them as evidence that the manuscript "contains meaningful text." The hypothesis has also shaped public understanding: Greshko (2025) writes that "paleographic analysis has also identified five scribes' hands within the VMS, suggestive of it being made by a community using a set of shared procedures" (p. 2).

The stakes are considerable. Davis has stated the interpretive logic explicitly: "If there were (at least) five people who knew how to write 'Voynichese,' it becomes more likely that the manuscript has underlying meaning, as opposed to a single author writing in gibberish or in a code that only they could read" (Davis, 2021a). The five-scribe hypothesis thus anchors a research program (Layfield et al., 2023) that includes hunting for semantic "cribs" beside illustrations, training AI on medieval manuscript images to find similar writing communities, and interpreting vocabulary differences between sections as evidence for distinct semantic topics. If the hypothesis is incorrect, these efforts rest on a false foundation.

Davis herself has articulated clear criteria for acceptable Voynich research: proposals must be "consistent with the realities of the object itself," must result from "a sound and explicable methodology that is logical and repeatable," and researchers "owe it to their readers, and themselves, to review what others have done (and cite them accordingly)" (Davis, 2020a).[1] This paper examines whether the five-scribe hypothesis meets these standards.

2. Currier's Observations and Their Reinterpretation

The five-scribe hypothesis builds on a tradition originating with Currier (1976), who observed substantial statistical differences between sections of the manuscript. Currier's actual assessment, however, was far more tentative than subsequent citations suggest. He described finding "not only two different hands" but "perhaps as many as eight or a dozen different identifiable hands," immediately adding: "Some of these distinctions may be illusory." Of the pharmaceutical section specifically, he wrote: "I cannot say with any degree of confidence that the last ten pages are in fact in the same hand as the first ten." His final estimate—"a total of something like five or six to seven or eight different identifiable hands"—is notable for its range and uncertainty. Currier himself characterized his terminology carefully: "my use of the word language is convenient, but it does not have the same connotations as it would have in normal use" (Currier, 1976).

This is a description of someone observing continuous variation and struggling to impose discrete categories on it. Currier saw differences everywhere, could not confidently draw boundaries, and arrived at a range so broad that it amounts to an admission that the variation does not partition unambiguously.

Davis's representation of Currier's work is markedly different. She describes him as having "quite correctly" discerned "two primary hands" in the botanical section, while characterizing his broader analysis as "incomplete, halfhearted, and somewhat unconvincing" (Davis, 2020c, p. 170). This converts Currier's honest uncertainty into two confident identifications and dismisses as halfhearted precisely the observations—widespread variation without clear boundaries—that most accurately describe the

[1] Davis further states: "To be taken seriously, proposed solutions must be subject to rigorous, expert pre-publication peer review" (Davis, 2020a, pp. 80).

manuscript's properties.

3. Hurych's Competing Analysis

A competing paleographic analysis has been consistently overlooked. Before examining it, however, it is worth noting that the single-scribe view has a long history. As early as 1946, Carter examined the physical manuscript and concluded: "The handwriting is incredibly consistent throughout. It varies only slightly in size, in slant, both of lines and of individual strokes, and in the flow of ink throughout the entire manuscript. It is the work of a single scribe" (Carter, 1946). Thirty years later, Currier disagreed—but, as shown above, with considerable uncertainty. Sixty years after Carter, Hurych (2007) investigated no fewer than 17 criteria—including slant, pressure, rhythm, margin, spaces, line drop, shakiness, upper and lower loops—and concluded: "We are convinced that the 'hands' A and B are the same, as well as the other 'sub-hands'. The differences observed simply do not justify the claim they belong to different persons."

Crucially, Hurych also observed: "While both 'hands' look visibly different, the idea that somebody would succeed in simulation of the first hand so closely as we can see in above table is rather remote"—in other words, if there truly were two scribes, one would have had to replicate the other's hand with a degree of precision that strains credulity.

Hurych further noted that "the writing B is flowing more easy or without longer pausing which was apparently necessary for sample A since it was written with more care," and offered a developmental explanation for the difference in neatness: "considering that sample A (folio 2) was at the very beginning of writing, the original care to put the text very neatly was probably lost when writing of the folio 26." This observation—decreasing care over time—is precisely what one would expect from a single scribe producing an extended text. It also suggests a developmental direction: the slower, more careful writing of Currier A is characteristic of a scribe who has not yet gained fluency in an unfamiliar script, while the faster, more flowing writing of Currier B reflects increasing skill and confidence. From this paleographic perspective, Currier A precedes Currier B—the same directionality independently established by the vocabulary evolution documented in Timm (2026a). This convergence of paleographic and statistical evidence strengthens the case for continuous development by a single hand.

Hurych's observations also align with Currier's own description of the pharmaceutical section as "cramped, slanted, having quite a different character, very obvious even to the untrained eye" (Currier, 1976)—a description of writing conditions, not different writers. Strikingly, Davis herself characterizes her Scribe 1 in identical terms: "This scribe's writing is broadly spaced and seems to have been written more slowly and carefully," while noting that Scribe 1's work includes "a large number of inconsistencies and 'mistakes' " whereas Scribe 2's work includes "relatively few of these inconsistencies" (Davis, 2021c). According to the five-scribe hypothesis, this contrast is puzzling—Davis cannot determine which scribe is more experienced. Following the single-scribe hypothesis, such behavior is expected: early in the process, the scribe writes slowly and carefully but makes frequent errors in a still unfamiliar script; later, fluency increases, writing accelerates, and inconsistencies diminish. To our knowledge, Hurych's work is not cited in any subsequent publication on Voynich paleography.

4. Methodological Concerns

Davis (2020c) emphasizes the benefits of the Archetype software system for digital paleography. However, the published screenshots suggest that only 44 of the manuscript's approximately 225 pages were uploaded for analysis. Davis (2022) confirms a reduced sample set: "Because the scribes on most of the first fifty-six leaves can be clearly identified as either Scribe 1 or Scribe 2, it was not necessary to upload all of those images for analysis" (Davis, 2022, p. 2). This statement is problematic in itself: it presupposes the result ("clearly identified") as justification for the sampling method.

More fundamentally, the diagnostic criteria for distinguishing scribes, the specific folios examined, and the process by which folios were assigned to scribes remain undocumented.[2] A telling admission appears on p. 176: "Folio 57v is somewhat problematic: there is too little text to reliably run Currier's dialect tests, and much of the text is composed of extremely rare characters, making a paleographical analysis difficult" (Davis, 2020c, p. 176). This sentence reveals that Currier's statistical tests are part of the scribe assignment process—when they cannot be run, the analysis becomes "problematic." If the paleographic identification were truly independent of statistical criteria, the inability to run dialect tests would be irrelevant. Without documented criteria, the five-scribe thesis is an untestable assertion.

The published paper's figures and Archetype screenshots consistently show only four scribal groups, not five. In a podcast interview, Davis confirmed that the fifth scribe was a late addition: "I initially thought there were four and then it was wait stop the presses there are five. I was checking my work. [...] So I had all the annotations for scribe three. I was looking at them and I kept seeing these outliers and then finally I just sort of embraced it" (Davis, 2021b, minute 15). Scribe 5, in other words, consists of outliers from Scribe 3 that Davis reclassified after finding them on a single conjoint bifolium. This is consistent with the observation that the published diagnostic criteria for Scribes 3 and 5 are nearly indistinguishable from those for Scribe 1.

5. The Diagnostic Criteria

5.1. The <k> glyph

A characteristic feature proposed by Davis to differentiate between two scribal groups (1,3,5 and 2,4) is the shape of the crossbar in the glyph $<k>$, specifically whether it appears bowed or horizontal. Davis explains: "Is the glyph formed by one or two strokes? Is the crossbar bowed, or is it horizontal? This is directly related to the previous question, since a bowed bar tends to result from a smooth directional change from the top of the first vertical, while a horizontal crossbar is the result of lifting the quill after completing the vertical" (Davis, 2020c, p. 172).

However, upon close examination, the variation in the writing of $<k>$ appears remarkably consistent throughout the manuscript. For nearly every page it is possible to find instances of $<k>$ written with bowed and with horizontal crossbar, respectively. Additionally, on nearly each page some instances of $<k>$ show an overlap between vertical stroke and crossbar, indicating two strokes regardless of crossbar shape. More strikingly, we observe also for the group of scribes 1,3,5 frequently cases where a small

[2] Davis has confirmed that her Archetype instance is no longer accessible: "I really cannot get at my bespoke version of Archetype anymore, but IT DOES NOT MATTER. [...] Archetype is a TOOL. It is not a calculator or a computer or a neural network" (forum post, September 29, 2025; `https://www.voynich.ninja/thread-4951-post-71295.html`).

foot is present at the base of the first vertical stroke of <*k*>. A foot is produced when the scribe lifts the quill from the writing surface, marking the end of a stroke. According to Davis, the base of this stroke should represent the starting point of the glyph, making the presence of such a foot incompatible with her single-stroke hypothesis: a starting point does not produce a foot, an endpoint does. Since handwriting inevitably produces variation in stroke shape, no scribe could consistently produce a perfectly horizontal or perfectly bowed crossbar. The observed variation is the normal fluctuation expected from any single hand.

5.2. The <n> glyph

Davis's second diagnostic feature is the "final flourish" of the glyph <*n*>. To our examination, this flourish shows a great variety of graphological details within a single folio. Pelling (2023) reaches the same conclusion: "From this page alone there would seem to be a wide range of scribal forms used when writing the terminating loop of the EVA 'n', ranging from really short to really quite long. Simply assuming that these 'can only be' scribal loops would therefore seem to be quite a foolish first step."

It is true that the final flourish of <*n*> is, on average, longer in Herbal A than in other sections. However, this difference has a straightforward explanation that does not require multiple scribes. The writing in Herbal A is broadly spaced and appears to have been executed more slowly and carefully—a difference known since Currier (1976), who described the contrasting pharmaceutical section as "cramped, slanted, having quite a different character." Hurych (2007) observed the same: "The writing B is flowing more easy or without longer pausing which was apparently necessary for sample A since it was written with more care." When a scribe writes slowly with ample space, strokes are longer and terminal flourishes more pronounced. When writing quickly in a cramped layout, strokes shorten and letterforms compress. The same hand produces visibly different results depending on writing pace and spatial constraints.

This point can be sharpened further. When comparing instances of the glyph group <*iin*> between Herbal A (e.g., f2r) and the Pharma section (e.g., f99r), consistent differences in the length of the final stroke are readily apparent. Since Davis attributes both sections to the same scribe (Scribe 1), these differences cannot reflect different hands—yet they are of the same kind and magnitude as the differences Davis uses to distinguish scribes elsewhere. The variation that is diagnostic between sections assigned to different scribes also occurs between sections assigned to the same scribe, undermining its use as a criterion for scribal distinction.

This internal inconsistency is difficult to reconcile with Davis's published methodology, which describes a focused process: "I initially annotated several different characters, but, after spending some time looking closely at different glyphs, I decided to focus initially on the single-loop gallows glyph" and presents the results as "preliminary" (Davis, 2020c, p. 168). In a subsequent response, Davis clarified that <*n*> and <*k*> are simply "useful diagnostics for quickly determining which scribe one is looking at," comparing them to diagnostic glyphs in known scripts such as *&* and *g* in 12th-century German manuscripts. Many other glyphs, she argues, "are paleographically useless" and describing them "would only add static and noise" (Davis, 2023). This analogy, however, is misleading. In known scripts, diagnostic glyphs serve as convenient shortcuts to scribe identifications that have been independently established through dated colophons, documented provenance, and identified scribal hands. For the Voynich Manuscript, no such external validation exists. The diagnostic glyphs

are not a shortcut to a conclusion established by other means—they are the only evidence. If they fail on examination, as shown above, nothing remains to support the identification.

5.3. The <sh> glyph

A further observation, overlooked by Davis, concerns the glyph <*sh*>. This glyph is formed by adding a plume stroke to the body of <*ch*>. (In some rare cases, a similar plume is also added to other glyphs.) In Currier A, the plume is for the majority of instances separated from the <*ch*> body, producing a visibly distinct form. In Currier B, by contrast, the plume is almost always connected to the <*ch*> body. If this difference reflected different scribes, it would constitute a genuine diagnostic criterion—a consistent, section-wide difference in how a specific glyph is formed.

However, the explanation is the same as for the <*n*> flourish. In the broadly spaced Herbal A pages, the vertical space available for each line of writing is greater. A separated plume requires more vertical room than a connected one. In the cramped layout characteristic of Currier B, where line height is reduced, there is simply not enough space for a separated plume, and the scribe attaches it directly to the glyph body. Once again, the same hand produces different results depending on spatial constraints. Both the longer <*n*> flourish and the separated <*sh*> plume in Currier A, and their compressed counterparts in Currier B, follow from a single cause: more space permits more elaborate strokes.

Notably, the two features do not change in unison. In the Pharma section written in Currier A, the flourish of <*n*> in <*in*> and <*iin*> is already shorter, while the plume for <*sh*> remains separated from the <*ch*> body. The flourish responds to spatial compression earlier than the plume, presumably because the two features have different sensitivity thresholds to line height reduction. This feature characterizes a gradual, continuous process—the same hand adapting incrementally to changing spatial constraints—rather than a discrete switch between scribes. It also means that had Davis selected <*sh*> rather than <*n*> as her diagnostic glyph, the plume separation would have tracked Currier A more consistently than the <*n*> flourish does. However, this would have produced a two-way split identical to Currier's original A/B distinction, not five scribes—and the spatial constraints explanation would have remained the same.

5.4. Statistical criteria masquerading as paleography

In her 2022 Malta Conference keynote, Davis introduced bigram statistics as scribe identification criteria: "René Zandbergen has recently observed that the work of Scribe 4 (Language B) can be defined by two additional tests: the relatively small frequency of the [qo] bigram, and the equally small frequency of [ed]. In other words, in addition to the shape of the [k], [n], and [f], the frequency of [qo] and [ed] can help identify the work of Scribe 4" (Davis, 2022, p. 6). Several problems attend this claim.

First, bigram frequency is a statistical property of the text, not a handwriting feature. A paleographer identifies scribes by how they form letters—stroke order, ductus, letterform proportions. Which character combinations appear frequently in the text has nothing to do with who held the pen. The incorporation of textual statistics into scribe identification conflates two independent analytical dimensions and creates the risk of a self-fulfilling prophecy: when topic modeling techniques subsequently display

correlations between scribes and text statistics, are they detecting genuine structure or merely echoing statistical input that previously entered the paleographic analysis?

Second, this is not a recent observation. Zandbergen wrote as early as 2016: "The very common character combination <*qo*> is almost completely absent in the zodiac pages and the rosettes page, but appears everywhere else" (Zandbergen, 2016). Moreover, Davis was aware of this observation before her first paper was published, having cited it in her 2020 presentation at the BSA Annual Meeting (Davis, 2020b, minute 31).

Third, Zandbergen's own conclusion contradicts Davis's interpretation. Zandbergen writes that "the overall statistics demonstrate that there is a continuum, and the other (not herbal) pages actually 'bridge the gap' " (Zandbergen, 2016)—evidence for continuous variation, not discrete scribes.

Fourth, the <*qo*> absence is not unique to Davis's Scribe 4 pages. Other folios not attributed to Scribe 4, such as the bifolio f1r, f1v, f8r, f8v, also exhibit exceptionally low <*qo*> frequency. The explanation is simpler than a scribal distinction: the Zodiac/Astronomy section contains exceptionally many labels, and it has long been known that the <*qo*> bigram is underrepresented in labels.

Fifth, the claim about "equally small frequency of [ed]" is factually incorrect. What Zandbergen actually writes is: "The very frequent character combination ed is almost entirely non-existent in all A-language pages" (Zandbergen, 2016)—a statement about Currier A, not about the Zodiac/Astronomy section. Davis appears to have attributed this observation to the wrong section of the manuscript.

The observations described in Sections 5.1–5.3 can be verified using an Archetype project containing annotated high-resolution images of all manuscript pages, available as supplementary material.[3]

6. The Five Scribes Reduce to Simpler Distinctions

A folio-by-folio comparison of Davis's scribe assignments with Currier's language classifications reveals a striking correspondence. Davis herself acknowledges the connection: "It was Currier who first determined that Scribe 1 writes in Dialect A and that Scribe 2 writes in Dialect B. The other three scribes I have identified—3, 4, and 5—also use Dialect B, at least according to the tests developed by Currier" (Davis, 2020c, p. 179). This statement retroactively projects Davis's scribe categories onto Currier's work—Currier identified dialect differences, not scribes—and is methodologically problematic: the Astronomical and Cosmological sections assigned to Scribe 4 were never straightforwardly classified as Currier B, since they consist predominantly of labels rather than running text and have always been treated as distinct from both A and B.

Nearly all Currier A pages are assigned to Davis's Scribe 1, and nearly all Currier B pages to Scribe 2. The mapping is near-perfect across the herbal sections (Quires 1–8) and extends throughout the manuscript. Davis claims to have identified scribes through paleographic analysis of handwriting features, while Currier identified languages through statistical analysis of character and word frequencies. These are ostensibly independent methodologies examining different properties of the text. Yet they produce nearly identical results.

There are two possible explanations. Either there really are discrete scribes who each used a distinct language—a remarkable coincidence requiring explanation—or Davis's

[3]The Archetype project is available at: Timm (2026b). doi: 10.5281/zenodo.18827108.

paleographic analysis was not fully independent of the Currier framework. Given that Currier's two-language distinction has been the dominant organizing framework for the manuscript since 1976, and given that Davis explicitly works within that framework (describing Currier as "quite correctly" identifying two hands), the simpler explanation is that the Currier A/B assignments shaped the scribe identifications. If one already knows which pages are "Language A" and which are "Language B," and one subconsciously expects handwriting differences between them, one will find them—because the handwriting does differ between sections, as a consequence of continuous evolution and varying writing conditions.

Davis's Scribe 4 is assigned exclusively to the astronomical and cosmological foldout pages. These pages are unique in the manuscript for containing mostly labels rather than running text. Labels are short, often single words or brief phrases placed beside illustration elements. They statistically behave differently from running text because they are not produced through continuous copying from adjacent lines. The statistical distinctiveness of these pages—including low $<qo>$ frequency—reflects text format, not scribal identity. This also explains why Sterneck et al. (2021) identified these pages clustering as a distinct "topic": any clustering method would separate label-heavy pages from running text.

In summary, Davis's two primary scribes replicate Currier's two languages under different labels, and her Scribe 4 identifies pages with labels rather than running text. The five-scribe hypothesis adds refinement (Scribes 3 and 5) to a framework that is itself a relabeling of pre-existing categories—categories that, as Timm (2026a) demonstrates, are coarse-grained abstractions of a continuous gradient rather than objective properties of the text.

7. The Circularity of Independent Confirmation

The five-scribe hypothesis has been presented as receiving independent confirmation from topic modeling (Sterneck et al., 2021), which in turn was endorsed in a linguistic review (Bowern and Lindemann, 2021), and further reinforced by statistical testing (Davis, 2022). Examination of the timeline, however, reveals that each link in this chain inherits from the previous one.

Davis was aware of Zandbergen's statistical observations about $<qo>$ distribution before publishing her paleographic analysis, having cited them in her 2020 BSA presentation (Davis, 2020b, minute 31). In her published paper, Davis writes: "I have sent my preliminary results to a professor of linguistics who is running several different linguistical analyses on the Voynich as part of a long-term class project and her own research. I have suggested that the work of the five scribes be analyzed separately to look for patterns that may distinguish them further" (Davis, 2020c, p. 177). Bowern—the unnamed "professor of linguistics"—then co-authored a study with her students that applied topic modeling to the text, using Davis's scribe assignments as one of its analytical categories. Sterneck et al. (2021) found that computationally derived topic clusters correlated with Davis's scribe assignments and concluded that this provides "strong evidence that the Voynich Manuscript contains meaningful text."

At no point in this chain is there a truly independent finding. Statistical observations shaped the paleographic analysis. The paleographic assignments were sent to the linguist with a specific analytical suggestion. The linguist's students performed topic modeling that recovered the same distinctions. Each step inherits from the previous one, and the original paleographic analysis was itself shaped by pre-existing statistical

frameworks.

Sterneck et al. (2021) conclude that scribes, NMF topics, and Currier languages "are not the same" and present this as evidence for independent semantic and linguistic structure. This is technically correct—three coarse-grained partitions of the same continuous variable will not perfectly overlap, because they partition at different granularities and use different features. But their imperfect overlap is not evidence of independent dimensions; it is the inevitable consequence of all three tracking temporal position in a continuously evolving process.

8. The Gallows Question

In a separate but related claim, Davis (2022) proposes that the single-leg gallows glyphs <*f*> and <*p*> function as abbreviations for the sequences <*te*> and <*ke*>, respectively. This proposal addresses a genuine observation: the double-leg gallows <*t*> and <*k*> are frequently followed by <*e*> (30.13% and 39.03% of the time), while the single-leg gallows <*f*> and <*p*> are almost never followed by <*e*> (0.72% and 0.36%).[4]

The abbreviation theory, however, introduces several problems rather than resolving existing ones.

(1) The proposed grouping—<*f*> with <*te*> and <*p*> with <*ke*>—is contradicted by contextual behavior. When a gallows glyph follows EVA-l, the most frequent forms are <*k*> (10%) and <*f*> (7.7%), while <*p*> (2.5%) and <*t*> (1.5%) occur less frequently. This groups the glyphs by leg count (<*k*>/<*f*> versus <*t*>/<*p*>), not by the abbreviation pairing Davis proposes (<*k*>/<*p*> versus <*t*>/<*f*>).

(2) Substituting rare glyphs with more common sequences increases overall text repetition, pushing the text further toward excessive redundancy.

(3) Words in the first line of a paragraph are already anomalously longer than those in subsequent lines (5.45 versus 5.07 characters on average); since <*f*> and <*p*> are overrepresented in first lines, the proposed substitution would exacerbate this discrepancy.

(4) The Voynich script already includes multiple sets of related glyphs with overlapping functions (e.g., <*i*>/<*ii*>/<*iii*>, <*e*>/<*ee*>/<*eee*>, <*ch*>/<*sh*>), suggesting that the gallows variants participate in a broader system of visually similar forms generated through the copying process rather than functioning as abbreviations.

9. Continuous Handwriting Evolution

Timm (2026a) demonstrated that the Voynich manuscript's vocabulary evolves continuously throughout the manuscript. The word network connects 84.67% of word types through single-glyph edits into a single structure (Timm and Schinner, 2020, Figure 2). Cosine similarity between folios "descends smoothly, al-

[4]Pelling (2020) proposed: "Might it be that EVA p and EVA f are nothing more complex than a way of writing EVA te and EVA ke?" Davis commented on Pelling's blog on April 5, 2022: "I hadn't seen this blogpost until today, and I'm glad I didn't see it before because I have independently developed a similar notion about p, f, te, and ke." Davis (2022) subsequently published a similar proposal without reference to Pelling's earlier publication.

most linearly" with no abrupt boundary between Currier A and B (Timm and Schinner, 2024, p. 308). Intermediate forms document gradual transformation: <*chol*> → <*cheol*> → <*cheo*> → <*chey*> → <*chedy*> (Timm, 2019). Vocabulary distribution is asymmetric: "words typical for Currier A exist in Currier B, but not the other way round" (Timm and Schinner, 2020, p. 8).

If the text was produced by a single scribe generating words through iterative copying and modification, the handwriting would naturally evolve in parallel with the vocabulary. As the scribe progresses through the manuscript over weeks or months, writing habits shift: stroke length, spacing, slant, and letterform proportions all change gradually. This is precisely what the manuscript shows. The visible differences between sections—broadly spaced, carefully executed Herbal A versus cramped, rapidly written Pharma and later sections—reflect changing writing conditions over an extended production period, not different individuals.

The interleaving of Currier A and B pages within the same quires further supports this interpretation. In Quire 4, for example, folios 26 and 31 (the two halves of one bifolio) are classified as Currier B, while the surrounding folios are Currier A. Under the five-scribe hypothesis, two scribes would have been alternating within a single quire, sometimes between pages of the same bifolio. This is implausible for a production workflow but entirely natural if one scribe batch-processed plant illustrations, writing some herbal pages early (Herbal A) and others later (Herbal B), with the bifolios from different stages becoming interleaved when the quire was assembled.

A difference "very obvious even to the untrained eye," as Currier (1976) described, is paradoxically more likely to reflect writing conditions than different scribes. Truly different hands, trained in the same scribal tradition, typically require expert analysis to distinguish—their writing looks similar precisely because they learned the same techniques. A difference obvious to anyone suggests a change in pace, space, or care, not a change in person. This is consistent with Hurych's (2007) comprehensive analysis and with the continuous evolution documented in the vocabulary.

10. Conclusion

The five-scribe hypothesis rests on undocumented methodology, diagnostic criteria that fail on empirical examination, statistical observations misidentified as paleographic findings, and a circular chain of seeming confirmation. The two primary diagnostic glyphs—<*k*> and <*n*>—show both variants on nearly every page of the manuscript. The statistical criteria (<*qo*> and <*ed*> bigram frequencies) reflect text format and evolutionary stage, not scribal identity. The five scribes reduce to pre-existing categories: Currier's two languages relabeled as two scribes, plus a label-versus-running-text distinction relabeled as a third scribe group.

The observable handwriting variation in the manuscript is real, as Currier recognized in 1976. But it is continuous—consistent with one scribe whose hand evolved alongside the vocabulary over an extended production period. Discrete scribe categories, like the discrete language categories examined in Timm (2026a), are imposed on a continuum. We echo Hurych's (2007) conclusion while remaining open to contradicting evidence—provided it meets the standards of reproducibility and documentation that scientific inquiry requires and that Davis herself has publicly advocated.

References

Bowern, C. L., and Lindemann, L. (2021). The linguistics of the Voynich manuscript. *Annual Review of Linguistics*, *7*, 285–308.

Carter, A. H. (1946). Some impressions of the Voynich Manuscript. `https://voynich.net/reeds/docs/carter.txt` (accessed February 28, 2026).

Currier, P. H. (1976). New research on the Voynich Manuscript. In M. E. D'Imperio (Ed.), *Proceedings of a Seminar*. Washington, DC: Privately Printed Pamphlet.

Davis, L. F. (2020a). What will it take to solve the Voynich Manuscript? *Manuscripts*, *72*(2), 73–85.

Davis, L. F. (2020b). Bibliography and Technologies: 2020 BSA Annual Meeting Panel Presentation. Video recording. `https://youtu.be/PdpekPqaB74`

Davis, L. F. (2020c). How many glyphs and how many scribes? Digital paleography and the Voynich Manuscript. *Journal for Manuscript Studies*, *5*(1), 164–180.

Davis, L. F. (2021a). The World's Most Mysterious Manuscript. *Medieval World*, *7*, 22–25.

Davis, L. F. (2021b). Scribes of the Voynich Manuscript. *Coding Codices*, Episode 2. Podcast. `https://soundcloud.com/coding-codices/ep-2-voynich-manuscript`

Davis, L. F. (2021c). Forum post. `https://www.voynich.ninja/thread-3188-post-45896.html#pid45896` (accessed February 28, 2026).

Davis, L. F. (2022). Voynich Paleography. International Conference on the Voynich Manuscript 2022, University of Malta. `https://ceur-ws.org/Vol-3313/keynote2.pdf`

Davis, L. F. (2023). Forum post. `https://www.voynich.ninja/thread-4035-post-55549.html#pid55549` (accessed February 2, 2026).

Greshko, M. A. (2025). The Naibbe cipher: a substitution cipher that encrypts Latin and Italian as Voynich Manuscript-like ciphertext. *Cryptologia*, 1–37.

Hurych, J. B. (2007). How many 'hands' wrote the VM? `https://www.as.up.krakow.pl/jvs/library/5-4-2007-09-06/` (accessed February 27, 2026).

Layfield, C., Zandbergen, R., Davis, L., Abela, J., Bowern, C., Rosner, M., and van der Plas, L. (2023). International Conference on the Voynich Manuscript 2022. *International Conference on Historical Cryptology*, 113–119. doi: 10.3384/ecp195700

Pelling, N. (2020). On single-leg gallows and the benefits of having more questions than answers. Cipher Mysteries blog. `https://ciphermysteries.com/2020/`

Pelling, N. (2023). Thoughts on daiin daiin. Cipher Mysteries blog. `https://ciphermysteries.com/2023/05/29/thoughts-on-daiin-daiin` (accessed February 26, 2026).

Sterneck, R., Polish, A., and Bowern, C. (2021). Topic Modeling in the Voynich Manuscript. arXiv:2107.02858 [cs.CL]. doi: 10.48550/arXiv.2107.02858

Timm, T. (2019). Shift from <*chol*>/<*chor*> to <*chedy*>/<*shedy*>. `https://tinyurl.com/chey-chedy`

Timm, T. (2026a). The Challenge of Analyzing a Dynamic Text: Why the Voynich Manuscript Resists Conventional Interpretation. LingBuzz. `https://lingbuzz.net/lingbuzz/009784`

Timm, T. (2026b). Archetype: Voynich Palaeography. Zenodo. doi: 10.5281/zenodo.18827108.

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

Timm, T., and Schinner, A. (2024). The Voynich manuscript: Discussion of text creation hypotheses. *Cryptologia*, *47*(4), 307–322. `https://doi.org/10.1080/01611194.2023.2225716`

Zandbergen, R. (2016). The Currier languages revisited. `https://www.voynich.nu/extra/curabcd.html` (accessed via Wayback Machine, March 18, 2016; current version dated 2019).