

James R. Child

## Again, the Voynich Manuscript

The Voynich manuscript has been called “the most mysterious manuscript in the world” with good reason. A host of professional scholars and amateurs have labored since at least the early part of the 17<sup>th</sup> century to make sense of the drawings of plants, astrological symbols, naked female figures emerging from tubes and, most of all, a unique writing system. The latter has been variously considered a cipher of a text (or texts) originally in Latin, Middle English, medieval Italian or French, or (more vaguely) “Teutonic.” The manuscript itself has been dated as early as the thirteenth and as late as the sixteen century with the majority view favoring late 15<sup>th</sup> or early 16<sup>th</sup>.

A number of scholars have dealt over the years with the physical description of the manuscript and the history of its provenience and ownership. One of the finest treatments in recent years is that of Mary D’Imperio, “The Voynich Manuscript – An Elegant Enigma.” Therefore, I refer the reader to that work and will not attempt a discussion of ink, calligraphy or other features already analyzed by professionals. I will instead attempt to assign consonantal and vocalic values to the twenty-odd symbols which I believe make up an alphabet. The authors of the manuscript apparently devised this alphabet (found nowhere else) to record information on a variety of subjects noted below. In a brief overview I follow the categorization developed by Professor William Newbold in his 1921 paper “The Cipher of Roger Bacon,” a classification system he apparently conceived from the kinds of drawings in the manuscript, as well as a (surely mistaken) notion that the accompanying texts discuss the drawings in enciphered Latin.

## Proposed transliteration with phonological notes, Voynich manuscript

Following this summation I will explain the rationale for my views regarding the nature of the alphabet and the language it reflects.

The manuscript contains the equivalent of 246 quarto pages. (Since eight leaves are missing it would originally have come to 262 pages.) Of these, the first and largest section has the equivalent of 130 pages, the vast majority of which (125) bear drawings of plants (with some truly strange features!) accompanied by text; that section is called the botanical division. It is followed by 26 drawings which are astronomical or astrological in nature but unlike any others known in those domains. Next comes a division of four pages of text and 28 of drawings (strange, almost alien in nature) of naked female figures emerging from tubes or pipes; these are referred to (for want of a better term) as biological (and called by one scholar “a plumber’s nightmare”). A fourth group contains 34 pages, again of flowers and plants but also including receptacles for processing medicinal substances, a set reasonably called pharmaceutical. Finally, there is a fifth division with 23 pages of almost solid text, which appear, in my research, to treat the phonology, morphology and syntax of the language of the manuscript itself, seemingly contrasting them with the corresponding features of kindred languages. The drawings in this section are few, being limited to the left side of the folia and serving in some cases as “paragraph” indicators. These pages I refer to as “linguistic”; they have proved to be the most accessible, although one or two of the botanicals have yielded valuable information, especially 23 recto which I *may discuss in another paper*.

I have researched the Voynich since the late 1970’s, following the procedures set forth below.

## Proposed transliteration with phonological notes, Voynich manuscript

My first challenge was to determine whether the symbols unique to the ms made up an alphabetic system (perhaps incorporating digits) or whether they were simply decorative flourishes expressing some esthetic purpose. While the latter was (and is) a possibility it has always seemed intuitively unlikely. Thus I have continued to concentrate on the former as the more promising approach: many of the symbols are similar to, or identical with the letters of the Latin alphabet, and some of those that differ are not greatly dissimilar to letter shapes of other familiar alphabets, such as the Cyrillic. I will therefore refer to all members of the set as letters (though it may turn out that some of them may double as letters and digits and a small number may stand for syllables).

To be sure, even if the symbols are mainly letters, it does not follow that they combine to form words producing what I will call “plain” (as opposed to enciphered) text. Their status as parts of meaningful segments of language (lexical or grammatical items) depends on where they fall relative to one another. To illustrate for English - the letters s-t-a-r form a high-frequency English word, s-t-a-m a potential one (e.g. an acronym of some kind which might become a word), while s-t-m-a is impossible in English as a word or a grammatical item; it could only be an abbreviation. These letter arrangements are either “canonic” (i.e., they represent actual or potential English words or grammatical elements), or they are “non-canonic”, as in stma.

English of course has many cases in which the writing system diverges widely from the spoken language. Thus, a cluster “ght” certainly does not represent today a pronounceable letter set, but it nonetheless occurs frequently in medial and final word position: “lighten”, “thought”, and so on. Similarly, French has two words beginning with the three vowel letters “oeu-“ (oeuf, “egg” and oeuvre “work”); otherwise the initial

## Proposed transliteration with phonological notes, Voynich manuscript

combination is very rare in that language. The two words (and their derivatives) are very frequent, however, so that from the point of view of the writing system the letter combinations are canonic. On the other hand, English words never end with “-hg”, nor do French words begin with “eou-”; such combinations suggest that any text containing them is either full of typographical errors or is an encryption of some kind.

To return now to the Voynich manuscript: do the symbols which are candidates for letters appear to group together repeatedly to form possible words or grammatical elements (such as verb tenses or noun plurals) in some “natural” language (and in a text written from left to right) or are they simply distributed at random? (The left-to-right direction seems guaranteed as discussed below on the section on paragraphing in 76 recto.)

As noted above, many of the symbols resemble those of known alphabets; thus, my next step has been to establish whether they distributed in patterns suggestive of plain text in a natural language (if symbols resembling, say, o and a bunch together in long strings, either the likenesses are coincidental, or the text is a cipher).

In fact, the symbols do seem to show patterns that point both to base forms of words as well as inflectional endings throughout the manuscript. However, in many of the divisions which contain great numbers of drawings the words appear in isolation, either as names of objects or their functions, or with some unknown purpose. Thus, it seemed most useful to concentrate on the “linguistic” division which not only yields word-like forms but also appears to link them syntactically: in English, for example, the isolated word “day” is canonic and very common. As an isolated part of a (presumably) deciphered message, however, it would count for little. On the other hand, if the

## Proposed transliteration with phonological notes, Voynich manuscript

decipherment included “The days are ? ing ?-er” the next segment might be “The days are growing longer”.

An emergent syntactic structure, with nouns and verbs seemingly reflecting agreement and government is most evident in the linguistic section (although it also is present in a more rudimentary form in the pharmaceutical and astrological sections). The patterning most resembles that of certain members of the Germanic family, for reasons discussed in the appropriate places. The folios I will give the most space in the discussion below are 76 recto and 106 verso, as these are the ones I have spent the most time on.

### Folio 76 recto

This folio has proved valuable in helping me infer a textual organization for the manuscript. First, it contains four paragraph-like segments in which the final line of each does not extend to the right-hand margin. Of special interest is the fact that the first of these – 29 lines long – is accompanied by nine symbols on the left-hand margin of the sheet which appear to be alphabetic: eight resembling consonant letters, one a vowel letter. Within these lines most of the letters are repeated in the textual “flow”, suggesting a linguistic treatment of the symbols in some kind of context.

*This needed  
or useful*

The other three paragraphs are much shorter, but, like the first, appear to treat linguistic topics. Each of the four is headed by what appears to be a particle resembling a preposition which is then conjoined with a longer string of letters, presumably forming nouns. I am of the view that these are introductory phrases (on the order of English “Regarding...”, “As for...” and the like) and that the author (or authors) intended to discuss the symbols mentioned above as letters representing the consonant and vowel sounds of at least one Germanic idiom. I have prepared for the reader a table of

## Proposed transliteration with phonological notes, Voynich manuscript

consonant and vowel letters I have assumed for the language, complete with Romanization (final page of paper).

### Paragraph 1

As noted above this paragraph, like the following three, begins with a string of letters the first two of which, , I posit as labial consonant such as f or v plus o. Taken together, this could be a preposition meaning "for" or "as to" and governing a noun, - perhaps a word having to do with a letter/sound, or even grammar part, in the language. Below I offer my transliteration of six "words" from Line 1 (of which the is the first); a copy of the pertinent part of the original folio is on a facing page.)

In addition, I offer a tentative English gloss for the words, including possible suffixes:

Fot<sup>1</sup>ot<sup>2</sup>ar <sup>3</sup>iefi<sup>4</sup>g<sup>5</sup>a g<sup>6</sup>oF<sup>7</sup> ...

For [?]s [?]ed from [roots?] with ...

1 2 3 4 5 6 7

The above is obviously not intended to yield textual information as such but it does offer a skeletal syntax several elements of which are reminiscent of certain Germanic languages. For example, in 2 and 3 there appears to be agreement between a plural noun and a past participle (cognate with English "-ed") and agreeing in number. In 4 a preposition common in Germanic, of, with a meaning of "from" seems to govern a plural noun in the accusative case (i.e. a direct object form in nasal a); this accords quite well

Proposed transliteration with phonological notes, Voynich manuscript

with Gothic and Old Norse parallels. 6 is perhaps one form a of a preposition (and prefix) very common in the manuscript: go meaning “with” or “through the agency of.” In the Germanic family the cognate forms ga-, ge- are usually verbal prefixes with perfective force, less commonly, markers for collective nouns: German machen “to make”, hat... gemacht “made”; halten “to hold, Gehalt “contents”.

There are many other segments on this page which suggest Germanic syntax; one, however, is especially intriguing because the formal pattern is fleshed out with interpretable word-forms. On line 14 beginning with word six (the word count is based on spacing) I cite the following (transliterated according to the accompanying chart):

gotar rīda rīda h̄ridā golīdā otiðā h  
1 2 3 4 5 6 7 8

This patterning, with two adjoining words identical in form and a third neighbor very similar in appearance, may be the main reason students of the manuscript have dismissed the possibility of its having been written in a natural language. Yet I believe I can make a case for a meaningful sentence:

“The Goths say rida [as] hrida, with the pronunciation of the preceding h.”

1 2 3 4 5 6 7 8

In 1 I have assumed the name of a people – the “Goths” – from the similarity of the base to the English word and the common North Germanic masculine plural -ar. (I must add

here that it is at the least premature to ascribe “Goths” to the ancient peoples of that name, or even to their descendants).<sup>1</sup>

Word 2 -- “say” – I have taken to be a cognate with a large set of Germanic words: German reden “to speak”, English “read” (and archaic rede “counsel”) and many others. (I should note here that the “roots” of content words cited above, and in other examples, are not restricted to North Germanic, as are the noun and verb suffixes, but may appear in any Germanic-speaking area).

3. This word, rida, is an excellent illustration of the writer’s practice of citing word-forms for their spelling/pronunciation. Identical in form with word 2, it is cited by way of contrast to word 4, hrida; it is as if we were to write in English “They read (aloud) [the word] read as hread.” To be sure this is an English sentence unlikely to see the light of day but it does bring out an interesting feature of the Germanic family whereby some languages have the initial consonant cluster hr- while some lose the h-: Gothic hrains, German rein “pure”. The writer makes this point in segments 5 through 8: g<sup>7</sup>o<sup>8</sup>t<sup>9</sup>i<sup>10</sup>d<sup>11</sup>a<sup>12</sup> “With the pronunciation of the initial h.” 5 (go-) I render “with” (discussed earlier); 6 I translate tentatively as “pronunciation” (one meaning of its Icelandic cognate); for 7, I ventured “preceding”, from a possible ot- “out, outside” (Gothic ut, us) and - i<sup>10</sup>d<sup>11</sup>a<sup>12</sup> “coming” (Gothic iddja). Finally, I assumed that 8, -h, refers to the h of 4.

The above is, of course, a tentative rendition. However, the folio has a number of syntactic frames similar to the above but without obvious Germanic cognates for the

---

<sup>1</sup> E. Prokosch discusses at some length the etymology of “Goths” in his work “A Comparative Germanic Grammar” published by the Linguistic Society of America, 1939, pp 29-30.

## Proposed transliteration with phonological notes, Voynich manuscript

content words; thus, an English “translation” would be too speculative. In any case I believe that the main point of interest is the author’s linguistic analysis (presumably written in his own language) of another Germanic language or dialect. The apparent dialectal difference regarding hr- versus r- I have already discussed but there are a number of others even on this single folio: the preposition/prefix go- “with”. For one, go- is the form native to the author’s language but he (she?) cites ga- (Line 9) as the form in a speech community —————→ which I have not been able to identify, even with reservations. (This prefix is, of course, the one used in fourth century Gothic, but there is no assurance that the same speech community is involved.)

The above discussion of 76 recto is obviously preliminary but further study may well throw light on many more text segments. I will now turn to 106 verso not only for additional examples of phonological comparison/contrast but also for some clues regarding possible patterns of stress or pitch.

### 106 Verso

This folio resembles 76 recto in that it is apparently organized to deal with topics embedded in paragraph-like sections, 14 in number as indicated by “stars” (but not letters) along the left margin. One of these – number seven from the top – is subdivided: the first part contains two lines in which the latter does not extend to the right margin; the second has four lines in which the last, again, is not full. (Immediately before the first word of the second subsection is a line slightly on the diagonal. This may be the indicator of a related but subordinate topic).

Proposed transliteration with phonological notes, Voynich manuscript

106 verso is intriguing in that it contains segments with Germanic cognates of content words operating with “little words” – prepositions, conjunctions, particles – to make plausible sentences. I will consider a few of these below.

Paragraph 1

The first line of this paragraph may raise questions of language history. I transcribe it:

Friða hoenā of erā ot'jodenā ...  
1 2 3 4 5 6 7 8

And suggest the following rendition:

“For words/speech –ing from early Jutish peoples...”

1 2 3 4 5 6 7

Paragraph 2, Lines 2-4 offer a more exploitable segment of text. It is in transliteration,

...oh tēnā ritar fōdā rītā riðā gotenā c. Tār aro  
soq fījōfīa dā ūnā frō tenā fōtēnā c. gotēnā fōdar  
ehðāðar horat or ēna līo ... A suggested translation:  
19 20 21 22 23

→ “... and those [?] writers[?] sound these [?] words [?] in the Gothic manner[.]  
These[?] are so raised<sup>(?)</sup> as to render frō tenā fōtēnā c.  
17 18 19 20 21 22 23

Gothic sounds are realized [?] higher than a/any form...”

Obviously there is a great deal of obscurity in the cited passage but the syntactic frame may well be valid. Thus words (2) and (3) appear to make up an adjective-noun phrase as a masculine plural subject – close to Old Norse. Likewise (5) and (6) have the

Proposed transliteration with phonological notes, Voynich manuscript

form of a direct object phrase (though the gender suffix for the plural noun form is closer to Old Saxon than Old Norse).

The two points of major interest in this passage, however, have to do with stress patterning on the one hand (14 and 15, 16 and 17) and the syntax of comparison on the other (items 21 and 22).

First, the question of stress. There is a good reason to suppose that in items 14 through 17 the writer is contrasting an initially stressed phrase with one having a final stress: *rotēna* vs. *rotēna*. (The original script includes the symbol *h* following the *o* in the first example whereas it has an *i* following the *e* of the second. These apparently are used as “lengtheners” for back vowels and front vowels respectively, as well as having the status of letters. See the transliteration table.) It seems likely that it is a question of the relative loudness of the stressed as against the unstressed syllable but it could also involve differences in pitch.

In regard to the segment which likely contains a comparison (*106 verso, Para 1 lines 3 and 4*): *gotēnā ḥodar ehṭādar horor or ēnā lijo ... (suggested translation: “Gothic sounds [are realized] more strongly than [any?] form ...”)*. This wording seems to indicate a more pronounced articulation (louder? higher?) for “Gothic” sounds than for certain others, the identification of which is obscure.

One final item of interest regarding 106 verso. In “paragraph” 9, Line 3, appears the clause oto lija godā oījā āsat [unknown digraph]. This segment

→ may perhaps be rendered as “according to the formula the good (saint? priest?) Ansgar.” This reference (if correct) would be to a Christian missionary to the northern peoples – Swedes, Danes and Slavs – in the mid- to late ninth century. Ansgar is referred to by contemporaries as an especially kind and cheerful saint/priest (the word

*oījā* → is very possibly a cognate of a set of Germanic

words meaning "holy", "sainted": Old High German wih "holy", Gothic weiha "priest", weihan "c<sup>anonize</sup>", (modern) German weihen "consecrate".

There are, however, some difficulties with the transliteration: two of the original symbols are quite rare. Still, if the suggestion should turn out to be right, it would provide a term inus ante quem: the authorship could not be earlier than the mid-ninth century.

It is clear from the discussion above that my judgments are far from certain on a great many points on the intentions of the author(s) of the Voynich manuscript as well as the symbolic system (i.e. assumed language) in which those intentions were embedded. I will end my treatment with as unbiased an overview as I can manage in the light of three decades of association.

First, I am confident that the manuscript is a text in a Germanic language, not, as some believe, that it is some kind of hoax or, as others would have it, an encryption of an underlying text in Latin or some other tongue. I have attempted to give support to this view in my paper.

Second, I am not sure about a few of the symbols to which I have assigned letter values. I believe most of my values are correct, otherwise I could never have developed a plausible (if unproved) phonological system supporting my thesis. Nonetheless, there are problematic gaps in the consonantal system with indications of a single symbol standing for a syllable, as well as two symbols representing a single value. I have included these in the transliteration table but a few comments are in order.

Surprisingly, the system does not seem to provide for an initial "n". Although my research points to the symbol g → as either ñ → (or an) → in medial or final position following a consonant, and as "-ñ" in final position following a vowel (one of the possible syllabic values mentioned above), I have not encountered a word plausibly beginning with "n". That is to say, I have found no potentially Germanic word cognate with English "night", "name", "near" or many another. This may be coincidental, though

the words cited above, as well as many others, are (obviously) very common in Germanic languages. On the other hand, the symbol  transliterated as "m" in the (assumed) word  "way, mode" may double as initial "n"; more instances of possible Germanic cognates in the manuscript are needed for even an educated guess in this question.

The outstanding case of two symbols for one value is the pair  for "e". Of great frequency, this digraph is found in words in which other (assumed) Germanic cognates show e. (Interestingly, a similar digraph appears in Gothic words with ai which also match e in most other Germanic languages).

The points above are, as noted, covered in tabular format. Here I will conclude with some general observations not amenable to treatment in tables.

To begin with, the Germanic languages in general are characterized by vowel changes referred to as "Umlaut" and "Ablaut". These terms, coined by German philologists, apply, respectively, to sound changes which are brought about either by assimilation e.g. the slightly archaic English plural "brethren" of (figurative) "brother", where the o of the latter word changes to e under the influence of the final e; or by grammatical change in which vowel alternations in the so-called "strong" verbs signal meaning change, as in English "sing", "sang", "sung". As far as I have been able to determine verb Ablaut does not play a role in the ~~+ex+~~ of the manuscript (although regular, or "weak" verbs show surprising affinities to the weak verbs of other Germanic languages).

In conclusion, the phonological and grammatical points I have raised above seem to show that the Voynich manuscript is in a natural language especially reminiscent of those of the Germanic family. Although I am far from certain about many points (and, obviously, at sea regarding precise lexical meanings) I feel that pursuing the approach I have used above (and with significant support from computer programs) will yield at least a partial solution to this centuries-old mystery.

Proposed transliteration with phonological notes, Voynich manuscript

Transliteration	Expansion	MS Letter Shape
a	e.g. English father	a
ã	nasal a	g
b	aspirated b/p	b
ð	e.g. English <u>the</u> , <u>then</u>	ð
e	e.g. English <u>bet</u>	e
ɛ	e.g. (approximate) English <u>they</u>	ɛ
g	initial position only	g
h	(1) like English consonant <u>h</u> before vowels (2) like the fricative <u>g</u> between vowels in some dialects of German (3) when final, lengthener of <u>a</u> and <u>o</u>	h
i/i	(1) as in English <u>pin</u> (2) consonantal <u>y</u> between vowels or between vowel and consonant	i
l̄	(1) palatal l (e.g. English <u>million</u> ) (2) li-	l̄
m	(possibly <u>n</u> as well, in initial position)	m
nã	word final	nã
o	(1) English <u>loss</u> (2) Possibly like English <u>w</u> before a vowel	o
p	unaspirated p/b (perhaps glottalized)	p
r	possibly "back" (French) or trilled (Spanish)	r
r'	palatalized r/ri	r'
s	Possibly English <u>s</u> or <u>sh</u>	s
t'	aspirated t/English <u>th</u> in <u>three</u>	t'
t	unaspirated (perhaps glottalized)	t
u	possibly like English <u>put</u>	u
?	unknown; rare letter or digraph	?
		CONSONANTS      VOWELS      UNKNOWN