

# The Cannabis Page of the Voynich Manuscript

J. Michael Herrmann  
University of Edinburgh

17th January 2018

## Abstract

The late medieval Voynich Manuscript (VM) has so far resisted decryption and has been even considered a meaningless hoax or an unsolvable cipher. Here, we investigate the hypothesis that the VM is written in a natural language by establishing a relation between the Voynichese text and classical Persian. In the VM not the standard Arabic-derived Persian alphabet is used, but a different alphabet that has similarities to Pahlavi and Mandaic script. Based on a transcription table, we can directly translate text from the VM which is shown for a page (f16r) that displays a cannabis plant. We obtain clear evidence for our hypothesis by the fact that a meaningful translation can be obtained for a continuous passage of VM text in a relatively easy and fully documented procedure. Although we do not attempt at a critical translation of the VM text here, we gain a first insight into the content of the VM, which may provide a basis for discussions of the provenance of the VM. The present attempt builds on earlier work of the author on the Pahlavi hypothesis, but revises some of the claims made in the previous paper [7].

## 1 Introduction

The Voynich Manuscript<sup>1</sup> (VM) is written on more than 200 vellum pages which date to the early 15th century (University of Arizona, 2011). Its history before the discovery by the bookseller Voynich in 1912 is largely unknown. Apart from a few cautious attempts, such as Refs. [2, 7], so far little progress has been achieved in deciphering the VM. There is not even an agreement, whether the VM has any meaningful content at all [11].

In a previous study [7], we have linked the VM to a Middle Iranian language, Pahlavi, which did exist around the time of the composition of the VM and still exists today. However, already in the 15th century, Pahlavi had ceased to be actively used outside the religious context of Zoroastrianism. We may thus conclude that either the VM was composed from Zoroastrian sacred literature (but may consist in parts of texts that are not conserved elsewhere) or other hypotheses should be taken into consideration. In the latter case and assuming that the Pahlavi hypothesis is “in the right ballpark”, a natural candidate would be classical New Persian. Persian was the dominant language in large parts of the Middle East including regions where Pahlavi was in use, such that a Persian influence is clearly possible. The Pahlavi hypothesis must be abandoned, because, as we will see, the VM vocabulary is surprising similar to the classical New Persian vocabulary that has changed only mildly for about a millennium. We should note, however, that, just as Pahlavi contains many heterograms, i.e. Aramaic loanwords that were pronounced as their Middle Persian counterparts, also classical New Persian contains many Arabic loanwords that have similar roots as the Pahlavi heterograms.

We will show that the VM is written in New Persian, although with an unusual alphabet. Voynichese letters are clearly different from the Arabic-derived standard Persian alphabet, while they can be related in part to the Pahlavi script and in part to the Mandaic alphabet which we will make evident here based on the identifiability of a meaningful Persian text from one of the VM pages. The Pahlavi hypothesis was thus successful only in so far as it turned out to be possible to use largely the same transliteration table (Tables 1 and 2, for changes compared to [7] see Sect. 2) in order to identify a word from the VM using a standard Persian dictionary [8, 13].

---

<sup>1</sup>MS 408 in the Beinecke Rare Book & Manuscript Library at Yale University



Figure 1: Page f16r of the Voynich manuscript. For a magnified view of the text, see Figs. 6 – 8.

After we present the transliteration table, we propose at a putative translation of the VM page f16r which features a picture that without reasonable doubt can be identified as cannabis, see Fig. 1. We find that the author of the VM does not dwell on the botanical or pharmaceutical properties, but mainly advises the reader to abstain from the use of the plant. In combination with the circumstantial evidence from a few other pages, the VM can be seen as an example of *wisdom literature*, although not as a particularly impressive one. The text largely appears to be, roughly speaking, a “rant” against any less wise fellows, rather than a sensible argument to convince feeble believers. We can speculate that the text has some gnostic influences (see Section 2.3). In the Persian-influenced regions as well as in Europe, the ideas of the Gnosis survived as an undercurrent in various mixtures with Christianity, Mandaicism, Zoroastrianism, Yazidism and Manichaeism. During centuries of suppression and persecution many of their scriptures were lost, such that the VM eventually could help to disentangle the complex interplay of these religions. A strong Islamic influence can be ruled out, because in this case it would be ununderstandable why the text was not written in the widely-used standard (Arabic) Persian alphabet. Instead, the similarity of the Voynichese alphabet with Pahlavi and Mandaic scripts can be taken to suggest a Zoroastrian or Mandaic background.

## 2 The Voynichese alphabet

Comparing the Voynichese (V) and Pahlavi and Mandaic scripts, we find that many V characters are rotated versions of letters from the other two scripts, see Table 1. This may be due to the different writing direction of the two scripts. A similar effect that was observed also in the earlier sinistrodextral Brahmi script [3], in which also some of the letters appear as upside-down adoptions from its likely predecessor Aramaic (right to left). Pahlavi, Mandaic as well as standard Persian (P) scripts derive ultimately also from the Aramaic alphabet. They have retained the dextrosinistral direction, while the VM is written in the opposite direction.

#	Voynich	Pahlavi	P/M	Mandaic	transcription
<i>o</i>		n.a.	→	◦	ā, ō, ū, [a]
<i>B</i>		𐭪	→	𐭫	b, [w]
<i>P</i>		◦	(→)	𐭬	p, f, [b]
<i>t</i>		𐭮	←	𐭭	t, d
<i>d</i>		𐭯 + 𐭮	←	𐭮	d, [z,j]
<i>r</i>		𐭰	←	𐭮	r, l
<i>cc</i>		𐭱	←	𐭮	s [ī, see Tab. 2]
<i>š</i>		𐭱	←	𐭮	š
<i>c</i>		𐭲	→	𐭮	ç, short vowel
<i>K</i>		𐭳	(↔)	𐭮	k, q
<i>g</i>		𐭰 or 𐭱	→	𐭮	q, k, g
<i>M</i>		𐭴	(↔)	𐭮	m [cmp. Tab. 2]
<i>8</i>		𐭵	(→)	𐭮	n
<i>w</i>		𐭶	(←)	𐭮	w
<i>a</i>		𐭷	←	n.a.	h, a, kh
<i>y</i>		𐭸	←	𐭮	y, i

Table 1: Voynich characters with the corresponding Pahlavi and Mandaic letters. The first column shows the V notation used here as a mnemonic code. The middle column gives a tentative pointer to the origin of the letter shape (either Pahlavi or Mandaic). See Notes below for comments. See also Fig. 2.

## 2.1 Comments on Table 1.

The letters are given in the order of the Persian alphabet for the main reading of each V letter. Frequently occurring deviations from the main transliteration are given in [brackets]. Strokes belonging to neighbouring characters were removed from the Voynich letters.

*o* One of the main differences to the Pahlavi hypothesis [7] is the identification of the letter V *o* with P  $\bar{a}$ , which is in most cases observed, or, more clearly, a Persian *'alef* is always represented by V *o* even it pronounced as a short vowel *a* or as a vowel different from *a*.

*B* As shown for V *K*, see Fig. 2a,b), there may be subtypes of the letter. It seems that sometimes the middle stroke leaves from the tip of the left stroke and sometimes it intersects the left stroke below the tip. However, we could not find enough evidence for the systematic use of these variants.

*P* This letters occurs often at the beginning of paragraphs, see Sect. 3. Again there may be subtypes of the letter that could distinguish between P *p* and P *f*.

*t* The visual similarity to Pahlavi *t* is only superficial as the strokes are different in both cases. The change of stroke order may be, however, due to convenience of writing after the letter was rotated. The letter has also similarities to some forms of Pahlavi *m*, see [4].

*d* The inverted breve diacritic that distinguishes this letter from V  $\check{s}$  is similar to the decoration that is used to discriminate Pahlavi *d* from *g* and *y* that are otherwise indistinguishable [9, Table II, p. xiii].

*r* Is very similar in shape with V *w*. We use as a criterion whether the lower end is parallel to the baseline (*w*) or whether it meets there diagonally (*r*). A kink is also evidence for the letter being a V *r*. In many cases both *r* and *w* need to be considered.

*cc* Denotes more often an  $\bar{i}$  than an *s*, see Table 2.

$\check{s}$  This letter is most clearly related to its Pahlavi original. It should nevertheless be noted that the Pahlavi letters (Psalter Pahlavi) shown in Table 1 are more than 500 years older than the Voynich text, such that the similarity may be simply a coincidence.

*c* This character is seen to reduce the ambiguity of abjad writing system by indicating the presence of a vowel, whereas pure abjad scripts write only consonants. V *c* is also used to represent a the *ayin* ( $\epsilon$ ) or the Mandaic *ušenna* symbol.

*K* See Fig. 2a,b).

*g* This letter occurs only in final position. See also Section 2.3, 2.

*M* This letter has two forms, see the discussion on letter V *m* below Table 2.

$\delta$  Represents the *n* sound. The symbol is chosen because of the similarity to the V letter shape. It is interesting that the Mandaic letter has a similar shape as the (inverted) Eastern Arabic number  $\blacktriangle$  ( $\delta$ ), while Voynich  $\delta$  is similar to the modern shape of the this figure.

*w* Neither Pahlavi nor Mandaic has a clear counterpart for this letter which is serves mainly the consonantic use of the letter *waw*.

*a* The Voynich shape is also similar to the shape of the final form of the modern Persian letter *he*.

*y* As in the VM, in Persian (but also in Aramaic) many words or grammatical appendages have an ending in *y*.











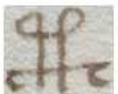







#	Voynich	Components	Transliteration
<i>aið</i>			<i>hā, ha, ah, a'</i>
<i>aið</i>			<i>hā', ha', a''</i>
<i>šB</i>			<i>šB, sB</i>
<i>šP</i>			<i>šP, šP, [šB]</i>
<i>šK</i>			<i>šK</i>
<i>šM</i>			<i>šM</i>
<i>m</i>		 (Pahlavi)	<i>m</i>
<i>cc</i>			<i>ī</i>
<i>co</i>			<i>ā, ah, a</i>

Table 2: Main ligatures and letter combinations from the VM. The first column refers to the uses the same V code as the first column in Table 1. The V ligatures are copied from f37r, the components in the second column are from Table 1. Strokes belonging to neighbouring characters were removed.

## 2.2 Comments on Table 2.

There is a similarity between the V ligatures and Avestan ligatures [1] which usually occur only for combinations *h* and *a* and of *š* with a following letter. In Avestan, however, *š* joins with *a*, *ce*, and *te*, while V *š* combines with other letters. It will be further interesting to compare the V ligatures with inscription of the much older Pahlavi coins [4, p. 89].

*aið* As in Persian, the combination like *ha* and *ah* are often interchangeable, such that it would be futile to aim at more specificity on the sequence of implied letters. There is, however, a correspondence between the stroke (including the final *ð*) and the number of elements from  $\{a, h, \varsigma\}$  with *ā* having two counts. There are also examples with a total of four strokes. Interestingly, the final up-stroke (*ð*) does not seem to have a literal meaning, which is in contrast to medieval Latin script where such sweeps can

represent an abbreviation. A similar combination occurs also with a final V *r* instead of the  $\partial$ , which then reads  $\bar{a}r$  etc.

*am∂* This ligature often represent the Persian plural ending.


*šB, šP, šK, šM*: Ligatures involving the letter V *š* (“table”) represent the succession of two consonants often in the beginning of a word, which does not exclude the occurrence of a short vowel between the two letters.

*m* One of the unsolved question (see below) is the occurrence of two different letter, V *M* (see Table 1) and V *m* for the *m* sound. While V *m* is quite clearly a rotated version of the Pahlavi character *mem*, the letter V *M* may have developed from Mandaic *m*. It is interesting that V *m* occurs only at initial position and is always joined to an *o* shape. Also *M* stands often in the beginning of a word, but (see e.g. the ligature *šM*) not exclusively. A hypothesis is the this ligature (similar to Pahlavi *vav-mem* that is affixed to ordinal numbers) denotes numbers, this is, however not the case on the cannabis page f16v. Another hypothesis relates V *M* to P *m* and V *m* to P  $\eta$ .

*cc* The ligature *cc* appears to be indistinguishable from the letter *cc* that is used for P *s*. Combinations of V *cc* and V  $\underline{cc}$  with V *c* can lead to an accumulation of several *c* shapes which introduce a high level of ambiguity.

*co* This may not actually be a ligature, but merely a script-style concatenation of V *c* and V *o*. Instead of being restricted to the implied transliteration to P  $\bar{c}a$  (or the unlikely combination of a P short vowel and P  $\bar{a}$  without an  $\bar{c}$  in between), it appears to denote similar combinations of P *h*, P *a* and P  $\bar{c}$  as the first ligature in Table 2, but within the word rather than at the end of it.



Figure 2: (a, b) Two versions of letter V *K* that appear to be represented in a contrastive setting on top of each other on page f66r. (c, d) In addition to Tables 1 and 2, a number of other letters occur in the VM (see e.g. f57v for more examples). The example c) corresponds possibly to P *j*, and the letter d) may be a P *z* (similar to Karoshti , *zain*, again rotated). Neither of the last two letters is present on f16r.

## 2.3 Examples

Before we present a translation of a full page, we will consider briefly a few examples from other parts of the VM that provide additional evidence for the VM being written in Persian. The examples will also give some preliminary insight into the nature of the content of the VM.

1. The VM starts (f1r) with the phrase *Ka šyr* which can be read as P *ka sher* meaning *like the lion*. This may be reminiscent of Ezekiel 32:2, but the text continues otherwise.
2. Page f65r contains only three words and a plant picture that has been identified as *Lady’s Mantle* (*Alchemilla*), but the describing text points rather to a flowering eggplant. The Voynich characters (see Fig. 3) are *oMang Nag atag*, for which we find quite unambiguously in Ref. [13]: *amhaq* (*of a dead white (like chalk)*), *nahq* (*parsley chiefly from fenny grounds*), and *hadaq* (*the brinjal or egg-plant*). It is not clear why *parsley* is mentioned here, perhaps the unusually lobed leaves hint at a species of egg-plant that is named in association with parsley. Since the three words all end in V *g* or P *q* they provide evidence for the reading of this rather infrequent letter in the VM.

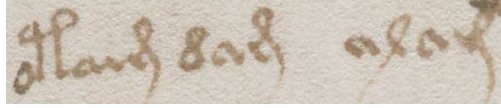


Figure 3: The text of VM f65r. The generous down-strokes (cut off in the figure) may indicate that not more text was planned here. For details see Section 2.3, 2.

3. The famous Rosetta page (Ros) of the VM has been interpreted in various ways. Just as a hint that a spiritual landscape rather than a geographical one is displayed, we translate the three pairs of words in the center of the middle circle on the right: (1) *oKavy oBataiθ*, (2) *oBššy Bary*, (3) *orauθ šy*, see Fig. 4. Ignoring the leading *V o* at the beginning of each of the pairs, we can transcribe this into (1) *qary* (*pursuing*) or *kahr* (*advancing*) or *qahrī* (*vindicating*) and *abad* (*eternity*), (2) *bašn* (*body*) and *bari* (*innocent*) or *bār* (*burden*), and (3) *rahā* (*liberation*) or (incl. leading *o*) *orāh* (*mind*) and *shai* (*wishing*). Inside the left upper panel on this page, we find only one word: *oBššarar*, which contains again *bašn* (*body*) and *alal* (*becoming decayed*). Even the T-O map in the upper right corner shows (again disregarding leading *V o* as “bullet points”) *bāl* (*heart, mind, soul*), *pārī* (*humility*) and *rad* (*intelligent*) rather than any reference to the known continents. Thus, even with some uncertainty of the translation, these statements cannot be seen as geographical information, but are clearly religious desiderata which may appear here as cornerstones of a theological or moral system represented on the fold-out page. To what extent this can be characterised as gnostic, remains to be discussed. Another fascinating speculation could arise, if it turns out the the VM is of Manichaean origin, namely, that the pictures are late copies from the lost Arzang (picture book) of Mani.



Figure 4: Text from the center of the middle disk on the right of the Rosetta page (Ros). The first two words appear bottom-to-top on left and the other two groups in one line on the right. See Sect. 2.3, 3 for details.

4. f67r1 shows a star-shaped center with twelve tips and 24 small areas surrounding the star. Twelve of these areas show each one or two stars, the other twelve are interleaved and represent a set of words. By translation to Persian, they can be identified as a list of twelve virtues, being described by the following adjectives: intelligent, splendid, simple minded, auspicious, blameless, modest, unbribable, prepared, inspiring, without wrath, good, penitent. The list has only partial overlap with the Aristotelian twelve virtues. The remaining text on this page seems similar in style to the example considered in Sect. 3.
5. f85r2 contains a prayer, see Fig. 5, that can be translated as follows

o bīr shir manār	O brother Sun, candelabrum,
shanī	salver of gold.
mānī dashn	Thou art an auspicious gift.
mīnā ār war āhār	Heaven, bring Thou our food.
t mair	Thou providest sustentation.
ār ār āk	Bring Thou, bring Thou wheat.

6. For a discussion of the “colophon” on final page of the VM, see Ref. [7] Appendix C.

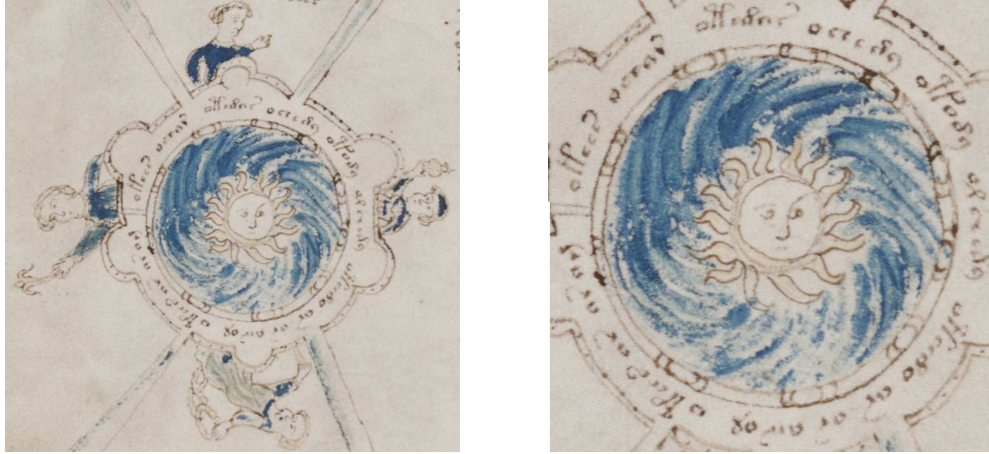


Figure 5: Center image of f85r2 in two levels of magnification. The image of the sun in the center is circumscribed by the text that is translated in Section 2.3, 5. The surrounding figures carry (east) a flask, (south) a chain, (west) a flower (?), and (north) a ring.

### 3 The Cannabis page (f16r)

The picture (Fig. 1) shows one of the few plants in the VM that can be identified unambiguously, see e.g. [12]. The expectation, however, that the text provides any information about its properties or uses, is soon betrayed by the “rant” the is accompanying the picture. The content of the text can be seen as reminiscent to Manichaeic or gnostic texts, but it is not the place here to attempt at a literary analysis.

Henning [6, p. 34] has argued that not *bang* or *mang* is the correct word for *hemp* in Pahlavi, but instead *šāhdānak* (*dānak* is translated as (*a small grain* [8, 9, 13])). This view has been challenged, see e.g. [5]. We note that *bang* or *mang* are not mentioned in the VM Cannabis text<sup>2</sup>, while *šahn* (*šan*) is used for *hemp* as also in the dictionaries [8, 9, 13], although it is not clear whether it is related to Pahlavi *šāh*.

We will first present a continuous translation of the text of f16r, see Fig. 1. The second subsection gives a word-by-word translation including alternatives for many of the words which are added in order to emphasise that the translation is by no means unique. Unless the text is translated using a correct methodology, it will certainly be possible not only to obtain a text of a different flavour<sup>3</sup>, but in many cases also a different partition into sentences or an essentially different content.

#### 3.1 A putative translation

##### 3.1.1 First paragraph

*Cannabis [is] vain. Stay away from the impudent crowing man. The pipe is a debasement. Jaundice [is] the overly happy face of the adherent. [He is] puffed up with pride. Security does not come [to him]. In the evening peace of mind does not come [to him, as] the serpent of nightly lust spoils him. [What is] concealed, will become public.*

##### 3.1.2 Second paragraph

*The debasement will be pointed to. Understanding does not help. Speaking obscenely even if astute is “gibberish”. The addicts make false promises. The meanness increases. Refrain from inquiring into the*

<sup>2</sup>The Pahlavi word *bang* translates as *henbane* [9]. It is the first word of f31r which shows *fleabane* according to Ref. [12].

<sup>3</sup>We should not expect a medieval text to comply with our current attitudes towards light drugs.



“gibberish” lies. [He] greedily eats like a pig. A receptacle boasting of of being filled. Those without knowledge, get ready! Stay away from stupefied sadness!

### 3.1.3 Third paragraph

The hempen rope is long. Speaking obscenely, he uncovers lamentable delusions of grandeur. Eternally [in the] fire you are! The herb is bringing ill luck. Smiling-faced [they are] desiring the deceit. Give him the pipe [he is] joyful and [has] a cheerful face, [but is] good for nothing, given up to destruction, being sick, [and] resembling grave diggers. His byname is pipe-bastard.

## 3.2 Details of the translation

The following tables present, paragraph by paragraph, all words from f16r together with their Persian and English counterparts. The Persian vocabulary can be inferred in most cases by looking up the Voynich letters in tables 1 and 2. While glottal stops, germination and vowel lengths (see notes below tables) are not precisely represented in the V script, the orthographic accuracy is quite impressive. Deviations from classical Persian orthography obey a regular pattern: Final V *t* corresponds often to P *d*. P *p* and P *f* are apparently not distinguished (see, however, notes to Table 1 and 1 in Section 3.2.1).

The English translation is mostly copied from the richer dictionary by Johnson [8] and the later dictionary by Steingass [13], both of which, in contrast to more recent dictionaries, include the vocabulary of classical Persian text such as the Shahnameh and are thus well suited for a text like the VM that is usually dated to the 15th century. Obviously, either is possible only because of the remarkable stability of the New Persian language throughout the last millennium.

The proposed translation is clearly far from being flawless. This is obvious already from number of alternatives that are added in the most ambiguous cases. For the translation of the text (last column), a choice was made based largely on context. Longer notes follow after each table. The tables are broken into single sentences if this is implied by larger inter-word spaces. The first column gives the line number  $\ell$  within each paragraph of f16r.

The word for which no direct Persian counterpart was found are V *MyBy* and V *NyBy* which form perhaps a play of words, see the comments below.

Persian is a *pro-drop* language, i.e. pronouns may be omitted when they are inferable. We find that the author of the VM drops pronouns frequently, such that the style become very brief, almost *telegraphese*, to a level that a sentences can consist of a single word only.

### 3.2.1 First paragraph

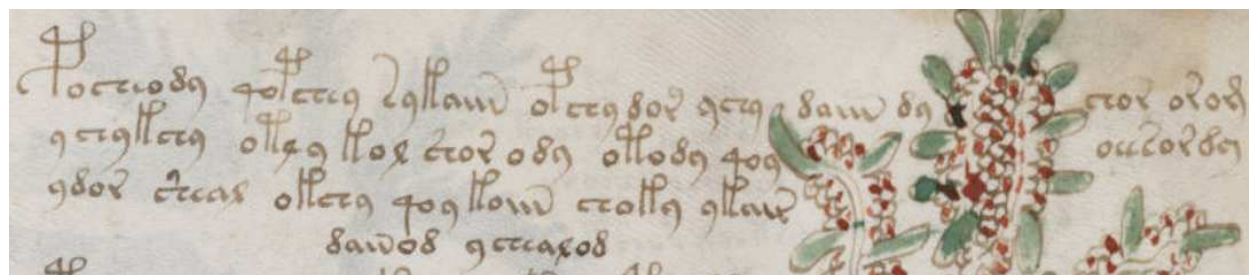


Figure 6: First paragraph of f16r.

$\ell$	Voynich	Persian	English [8, 13]	notes	translation
1	P	f[as̄l]	section, chapter, article	(1)	—
	o		[sentence initial]		—
	šco8y	šan	amorous blandishments; hemp	(2)	Cannabis'
	moPšcy	mafšūš	emptied, voided; idle, vain	only in [13]	vanity,
		mufaššl	wind-expelling	(3, 4)	
		ne + pushi	not + put on, wear	other verbs?	
	wyBanð	waib	fy upon, woe to (you, ac.)	plural	woes!
	oPšy	afšādan	to strain, squeeze, press; to speak obscenely, profanely, boorishly		The impudent
		afšār	speaking idly or obscenely	in compositions	
	8or	nar	male	connected to previous	man,
	yšy	yašš	glad, rejoicing	(5)	crowing,
	8anð	nāh	refrain from		abstain from!
	8y	nai, nay	pipe, tube, flute, fife; a reed, cane, pen; the windpipe, throat, gullet		The pipe [is]
		nai, ne	no, not, neither, nor		
	šor	šār	adulteration, debasement		a debasement
	orog	irāqat	spilling, shedding, pouring forth		
		urāq	jaundice		Jaundice
2	yšy	yašš	glad, rejoicing,		
	Bšy	bašš	of a cheerful, open countenance	connected to previous (6)	[is] the overly happy face
	oMty	ummatī	a follower of a religious creed	letter V <i>t</i> has unusual shape	of the adherent,
	Bot dor	bād-dār	windy; swollen; puffed up with pride; independent; worldly-minded; possessed of an evil spirit; destructive	(6)	[he is a] puffed-up-with-pride
	o8y	ānī	that which has an individual existence		person
	oMo8y	amānī	security; trust, charge, deposit	(7)	Security
		amām	before, prior to		
	moy	māy	any stinging or biting animal, as serpents, leeches, etc.; do not come	(8)	does not come [to him].
		ma'y	extending, stretching	(9)	
	occwor8y	ewār	evening	$\underline{cc} = \bar{i}$	In the evening
		aswārī	a mode of warfare with hatchets and maces	$\underline{cc} = s$	
3	y8or	nār	fire; hell; a mark made by burning; mind, intellect; counsel, advice		mind
		y nar	that man		
	dcat	da'at	quiet, tranquillity; easy circumstances		tranquillity
		du'āt	invocations, prayers		
		da'ṭ	a driving, a thrusting		
		da'd	name of a woman celebrated for beauty		
	oBšy	abašš	smiling-faced	[8]	[to the] smiling-faced

		ābiš	Abyssinia		
		abš	collecting		
		biš	give to him		
		bašš	of a cheerful, open countenance		
		baš	a bandage; a hoop of silver, iron, or brass to fasten boxes; a padlock; a watered field		
		buš	a horse's mane; a fringe, flounce; a lock, ringlet; incomplete		
moy	māy		any stinging or biting animal, as serpents, leeches, etc.; do not come	connected to next? (8)	does not come
	ma'y		extending, stretching		
Bonθ	bāh		lust; sperm; sexual intercourse		[as] the lust
	bāh		soup, broth, meat, victuals; be it so	for <i>pāh</i> or <i>bād</i> , resp.	
	bāhā		a kind of food		
	ba		good, excellent; elegant; better; safe, sound; quince		
	buh		owl		
šoBy	šabī		nightly		nightly
	šāwidan		to become; to go	2. pers. present	
yBair	yaba		injury, mischief; spoil, damage	(10)	spoils [him].
	bahār,		spring, beginning of summer; a blossom; orange-flower; a Buddhist temple; an idol; the harem of a prince; camomile, anthemis; name of an island; a species of odoriferous herb; bupthalmus, ox-eye; anything beautiful and splendid		
	buhār		idol		
	bahr		shining, excelling (in beauty, grace, virtue); excellence, preeminence; conquest; the imposition of a burden beyond the strength; astonishment, wonder		
	buhr		shortness of breath, asthma, breathing hard from overload; a city, town, or province; the middle of a valley, a channel		
4	𐭪𐭣𐭥𐭥	nihān	hid, concealed, occult; prudently, secretly; the mind, soul, interior		[What is] concealed,
	yšcato8	išā'at	publishing, divulging, diffusing, spreading about or abroad	compare Mark 4:22	becomes public.
		yaštan	to pray in a low murmuring tone as the fire-worshippers do at meal-time; to beg, supplicate; to celebrate		
		šaiṭān	Satan, the devil		

## Notes

1. Many of the pages and in some parts of the VM also paragraphs, start with a letter *V P*. Whether or not it actually represents the indicated Persian word for chapter etc. cannot be decided here. Because it does not represent the first letter of a meaningful Persian word in most cases, we can conclude that initial *P* indeed functions as a paragraph marker. Although this was not further checked here, the letter *V P* may actually have two variants that differ in the curl of the left-ward stroke and which would naturally represent either *P p* or *P f*.
2. Considering also the trailing *V o*, a number of other translations of the first word (e.g. related to drinking: *āshānīdan*) are possible. They are not considered here based on the evidence from the image on this page.
3. This is considered as an unlikely side effect of cannabis consumption.
4. The *-ing* forms in the English vocabulary are all verbal nouns.
5. The indication of long consonants (germination) by *shadda* is generally ignored in Voynichese.
6. An *alliterative or rhyming combination* [10, p. 68].
7. Initial *V o* often represent *P a* rather than *P ā*, while within words it is usually *P ā*. In some cases, leading *V o* is a word or sentence separator, analogous to a bullet point.
8. As it does not contain a labial, the *V* word *moy* poses a problem to our explanation of the occurrence of two letters (*V m* and *V M*) for the *m* sound. The word occurs twice in this page.
9. The *ayin* (◌) is sometimes represented by *V c* or *V ɪ* but more often ignored. This custom is akin to the sloppy use of ◌ in Arabic (loan) words in Persian.
10. The last letter has the shape of an *r*, but may actually be *ɪð* due to lack of space close to the plant drawing.

### 3.2.2 Second paragraph

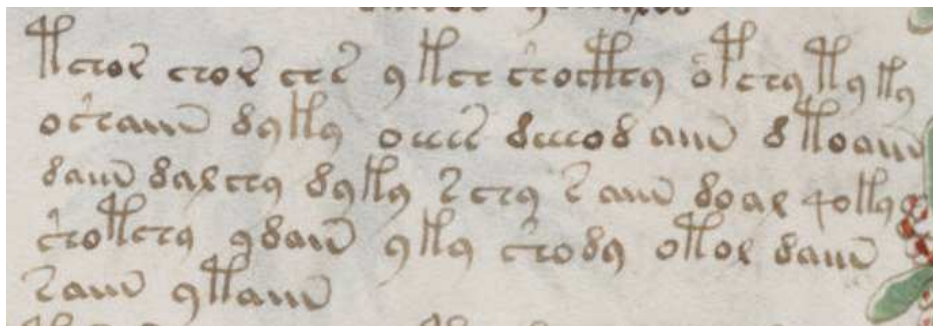


Figure 7: Second paragraph of f16r.

$\ell$	Voynich	Persian	English [8, 13]	notes	translation
1	Mšor	mušār	signified, indicated		Indicated
		mašār	hive of bees		
	šor	šār	adulteration, debasement		the debasement.

	šw	šav, šū	be, exist; be thou; (in comp.) being, existing	<i>šav = šab</i>	
		šav	night		
	yBš	yābiš	comprehension, understanding, intellect, knowledge	V o absent	Understanding
		yab	pointed arrow		
		bašš	of a cheerful, open countenance		
	došMy	dušmat	worthless, good for nothing		does not help.
		duš	evil		
		do	two		
		šam	terror, flight, ... (see above)		
	oPšy	afšādan	to strain, squeeze, press; To speak obscenely, profanely, boorishly		Speaking obscenely
		paš	the mane of a horse; a fore-lock; a tuft hanging loose at the end of a turban, sash, or belt; ignoble; defective; like, resembling		
		piš	before	for <i>peš</i>	
		puš	owl		
		piši	a small, thin, copper coin		
		pašij	Circular (as a piece of money), fish-scale, etc.); abundant, ample; many; apparatus for a journey		
		pašijidan	to be armed; to prepare		
	My_By	may	wine (etc.)	(3)	“even if”
		mī	imperfect tense		
		maibid	and		
		mai-bih	a kind of medicine made of quince-juice		
		bī	without, in-, un- etc.		
2	odanθ	adhá	more or most subtle or astute		astute
	8y By	?			[is] “gibberish”.
		naib	striking on the canine-tooth; breeding a canine-tooth	(3)	
		nib	a particle sometimes added to words without altering the sense		
		niwe	writing, scripture, book	Pahlavi: <i>nibēg</i> [9]	
		ne bī	not without		
	occcw	asīr	bound; prisoner, captive, bondman	(1)	The addict
	8cc_co8	nisān	discord, opposition; a false promise, a lie; the time when a promise is broken		makes false promises.
		naisan	like, resembling		
	anθ	ha˘	greedy; covetousness, avarice, meanness, baseness; cowardice, timidity; take! give! grant!		[His] meanness
		hā	(the 3rd pers. sing, of the affixed pronoun) her, to, of, from, or by her; (referring to broken plurals) theirs, etc.; (interjection) lo! behold! take!		

	8Moanð	namā <sup>c</sup>	growing, increasing; rising (water); rising (as the price of provisions); interests (of a capital)		increases.
		namā, numā	showing, pointing out; an index		
		na- mānad	may he (she, it) not remain, i.e. may he perish!	<i>na-manad</i> is also used	
3	8auð	nāh	refrain from		Refrain from
	8atšy	nadš, nadaš	inquiring into, investigating; dividing, separating, carding (cotton)		inquiring into
	8yBy	?		(3)	“gibberish”
		naib	striking on the canine-tooth; breeding a canine-tooth		
		nib	a particle sometimes added to words without altering the sense		
		ne + bī	not without		
	wšy	wašy	painting, printing, staining, dyeing (cloth or stuff); lying, colouring a story with falsehoods; reporting to the king, informing against; the colour or painting of a garment; embroidery		lying.
	r	r	if	connected to next?	If
		ra <sup>cc</sup>	tranquillity, quiet, rest		
		rah	see thou, the mercy of God be upon him		
		rahh	wide and shallow (dish)		
		raha	liberation, escape; liberated		
		rahā <sup>c</sup>	wide; a large district		
		rihas	drawing near, approaching; consenting		
	auð	ha <sup>c</sup>	greedy; covetousness, avarice, meanness, baseness; cowardice, timidity; take! give! grant!		[he is] greedy
		hā	(the 3rd pers. sing. of the affixed pronoun) her, to, of, from, or by her; (referring to broken plurals) theirs, etc.; (interjection) lo! behold! take! (particle of swearing) by		
	8oat	nā <sup>c</sup> it	one who eats ungracefully, who swallows one half of a mouthful, and lets the other drop on his plate, or throws it away; one who travels to a great distance		[he] eats like a pig.
	4oBy	nabīy, nabī	prophet	Pahlavi: <i>nabī</i>	
		nubīy, nubī	high grounds		
		nāb	pure, clear, limpid, genuine, unadulterated		
		ma’abī	pestiferous		

		ma'āb	returning; a place of return; a receptacle, repository, a place where anything centres, or in which it is contained		A receptable
4	doMšy	damiš	breathing, blowing; assault; boast; eruption, efflorescence; the marrow of a tree; the middle of anything		boasting of
		dām do	snare, trap, gin, or net; vanities and deceptions of the world, worldly illusions (metaphorical)		
		dāmi <sup>c</sup>	Shedding tears, tearful		
		dāmi <sup>c</sup> at	(wound or fracture in the skull) letting blood flow		
	y8aiθ	nah	being filled, replenished; being half-raw, ill-dressed (meat)	y?	being filled.
		īnah	In, this →īnha, those		
	yBy	y + bī	who without		Those without
		yab	pointed arrow		[or: the arrow of]
	do8y	dān	grain; (an affix denoting) what holds or contains anything; (imp. of dānistan) know thou; (in comp.) knowing, intelligent		knowledge
		dānī	approaching; near, within reach; low, rile		
	oMot	āmadan	to prepare, make ready, equip, expedite, despatch; to be prepared; to fill		prepare.
	8aiθ	nāh	refrain from		Refrain from
5	waiθ	wahh	grief, sadness	[8]	sadness
		wah	Strange! would to God! alas! have a care!	(2)	
		wāh, wah-an	(an interjection of admiration) O! wonderful! excellent! (also an interjection of pain) oh! alas!		
		wā-hā	meats, dishes		
	yMaiθ	āmāhidan	to swell, to be stupefied		stupefied.
		mahā			
		māh	moon, fish etc.		

1. The third V *c* is not connected, while the previous two are. An exact match would imply the V expression occ-ccr. In the present spelling the second vowel would be short.
2. While the VM is generally very clearly written, it is sometimes difficult to distinguish the letters V *r* and V *w*, i.e. alternatives such as *raviθ* or *raviθ* (see also line 3 in this paragraph) cannot immediately be ruled out.
3. For V *MyBy* no simple translation could be found. We assume that it forms together with V *8yBy* a wordplay which may be analogous to *pribble-prabble*, *claptrap* or *flubdub* in English. The part *8yBy* occurs in the next line in isolation, so we can assume that the second component contains the lexical unit (as in English *chitchat*). That *nibi* means *dialect* in Hebrew may be a mere coincidence. The translation by *even if* and *gibberish* is merely a (partially self-describing) placeholder.
4. It is not uncommon in Persian that *yam* is put for *am* after nouns ending in *a* or *o*.

### 3.2.3 Third paragraph

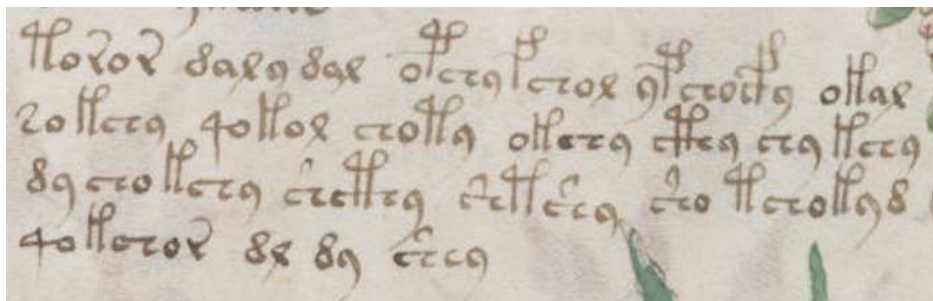


Figure 8: Third paragraph of f16r.

$\ell$	Voynich	Persian	English [8, 13]	notes	translation
1	Moror	marār	a hempen rope	[13]	The hempen rope
		mirār	being drawn out; wrestling, striving to throw another on the ground; going along with, accompanying, times, turns	[8]	
	Ṣaty Ṣat	natnat, natānit	tall, long		[is] long.
		nad	increase, abundance; bent; good guidance		
		nadd	fleeing, running away, scampering off; a perfume composed of musk, ambergris, and the wood of aloes; or simply of ambergris, aloeswood	in P. also <i>nad</i>	
		nidd	image, idol; like, resembling; an equal, peer, friend; an enemy, rival; scented with ambergris		
	oPšy	afšādan	to strain, squeeze, press; to speak obscenely, profanely, boorishly		Speaking obscenely,
		afšū	come	[8]	
		ifšāc	divulging, revealing, disclosing, publishing; publication		
		piši	a small, thin, copper coin.		
	Kšot	kušād	he opened, displayed, uncovered; he exhilarated; he seized; an opening, uncovering; disclosure, revelation; cheerfulness; expansion; a siege; the shooting of an arrow; a move at chess; open, wide		he uncovered
	yPšošKy	shakkī	doubtful; habitually complaining, querulous; lamentable; indisposed, slightly ailing.		lamentable delusions of grandeur.
		ibšāš	having luxuriant, entangled herbage (ground); putting forth its early grass	<i>aky</i> as a diminutive ending	



		ifsha <sup>c</sup>	being proud, arrogant, playing the great man [8]; divulging, revealing, disclosing, publishing; publication		
		yaf	ascending (a mountain); approaching twenty years of age (a youth)	(1)	
		yafa <sup>c</sup>	hill; tall, well-grown (youth).		
		šašak	a four-stringed instrument; name of a note in music; a bird resembling a partridge, but smaller; an animal resembling an ape		
	oBat	abad	eternity without end (opposed to azal, eternity without beginning), perpetuity; of one year's age or growth	puts forward in a year,	Eternal[ly]
		ābad	perpetuity; a wild untractable disposition; angry, affrighted, scared; the eye		
		ibat	being contracted; blushing; disgrace, shame, ignominy		
2	wo	wā	fire; distant; with; again (as prefix), woe!	unconnected to next	[in the] fire
	Bšy	bāši	you are (bāšīdan)		you are!
		baši <sup>c</sup>	stinking-breathed		
	4oBot	nabāt	plant, herb, vegetable; vegetation; fine sugar, white, candied, refined		The herb
	šoMy	ša <sup>c</sup> m	being unlucky; bringing ill luck; black moles or marks		brings ill luck.
		šamm	smelling		
	oBšy	abašš	smiling-faced	compare <i>bašš</i> above	Smiling-faced
		ābiš	Abyssinia		
		abš	collecting		
	šMy	šam	terror, flight; the tail; deceit, fraud, trick; distance; imp. of šamldan; (in comp.) terrified, astonished, fearing; disturbed, distracted		the deceit
	šy	šay <sup>c</sup>	wishing, willing, desiring	(2)	[he is] desiring.
	Bšy	biš	give to him	for <i>bi-dihaš</i>	Give to him
		bašš	of a cheerful, open countenance		
3	8y	nai, nay	pipe		the pipe [he is]
		nay, ne	no, not		
	šo	šā	joyful, glad; a fruit-tree	for šad [13]	joyful and
	Bšy	bašš	of a cheerful, open countenance		[has] a cheerful face,
		biš	give to him	for <i>bi-dihaš</i>	
	dšMy	dušmat	worthless, good for nothing		[but is] good for nothing.
		dišum	afternoon, evening prayer		
		duš	evil, bad, ugly		
	dMdy	damdamat	given up to destruction		[He is] given up to destruction.
	do	dā <sup>c</sup>	being sick; disease, sickness		[He is] is sick,

		do	two		
	MšoByn	mušābih	resembling		[and] resembles
4	4oBšor	nabšar	grave-digger		a grave digger.
	8t	naˁt	epithet, description; an adjective noun; praise, eulogium, encomium	see <i>nad</i> above	[His] byname [is]
	8y	nai, nay	pipe		the pipe
		nayˁ	ill-dressed, half-raw	(3)	[or: a half-baked]
		nay, ne	no, not, neither, nor		
	dcy	daˁiy	a spurious child, bastard; an adopted son; one about whose pedigree suspicion entertained	[8]	bastard.

1. In the VM text, a part of the *š* is missing in the *šK* ligature. This may be intentional as this construction occurs, although rarely, also in other pages of the VM.
2. There is no larger inter-word space here which may be due to the lack of space near the drawing.
3. The identical spelling of *8y* should imply the translation *pipe* as above. The possibly less likely translation was chosen here because the author may have attempted to use a more dramatic wording in finishing the “rant”.

## 4 Discussion

The present approach does answer a few questions that have been mentioned before in the literature.

- The VM can be read as a natural language. This high level of orthographic coincidence with Persian words provides substantial evidence for this hypothesis.
- Persian is a pro-drop language, i.e. sentence constructions are possible that are very short. The mannerism of the author may have contributed, however, to the fact that as a text the VM is still hard to read.
- The frequent occurrence of *y* endings in the VM is also common in Persian.
- The frequent occurrence of the final group  $\bar{w}d$  can be explained by reference to the Persian plural ending. Obviously, this ending occurs also in other cases.

Remaining questions within the current approach include

- Why are there two letters (denoted here as  $V m$  and  $V M$ ) representing for the *m* phoneme? Although  $V m$  tends to occur often in front of labials (*b*, *p*), this is not exclusively so, as seen already by two examples on f16r. Therefore the hypothesis that  $V m$  represents the P letter *n* if it is pronounced as *m* in front of labials (sometimes denoted by  $\bar{m}$  in Persian), is insufficient. It may be possible to study parallels to Avestan and various Indian languages, where more than one *m* sound exists, in order to disambiguate the two versions *m* and *M*.
- Why do certain letters occur in certain places only? E.g. letter  $V g$  occurs only in final position, while the “capitals” are frequently in front of a word. The inhomogeneity of letter position probabilities this may be a feature of the New Persian language and related to the use of initial, medial and final letter forms in (Arabic) Persian script. This may imply that additional transcription variants will turn up. E.g. a P *g* or P *gh* in a non-final position may be represented by a  $V K$ . An *M* in final position will probably appear as an *n*, i.e.  $V \delta$ , as it is also possible in standard Persian.

- Why the style of manuscript telegraphese? Where is more precisely the origin of the writer, what is her or his cultural or religious background?
- Why does the VM include illustration if it does not provide pharmaceutical, biological or medical information? The quality of the illustration is admittedly no match to the state-of-the-art science at the end of the byzantine era, so a scientific purpose is apriorily unlikely, but a Herbal in a folk-medicine context may be fine with pictures of this quality. It is of course possible that other pages provide more detail related to herbal medicine, while the authors dissented to disclose information on the positive effects of cannabis.

These questions are not meant to foreclose other important questions that have been asked by other VM researcher and which may turn out to be more relevant towards a actual translation of the manuscript. The main question to be asked is, why a plain Persian text should have escaped the attention of many researchers. We should return to this question when there is more evidence for the present approach.

As already initially stated, we did not attempt at an philologically acceptable translation of the even just a single page, but to provide evidence for a hypothesis that may be sufficient to stimulate more interest within the Iranian studies. At least we can say, that the proposed approach has the advantage over other solution attempts that it is easily falsifiable as anyone in command of classical Persian language will be able to judge whether (apart from a most likely substantial number of minor errors) the presented hypothesis is invalid. If this is not so, then we can be confident that a complete and critical translation of the VM will be soon available. Otherwise, the adventure will continue.

## References

- [1] Avestan. *Chapter 10.7, The Unicode Standard, Version 10.0*. Unicode, Inc., Mountain View, CA, 2017.
- [2] Stephen Bax. A proposed partial decoding of the Voynich script, 2014.
- [3] Georg Bühler. *On the Origin of the Indian Brahma Alphabet*. K. J. Trübner, 1898.
- [4] Carl Faulmann. *Das Buch der Schrift enthaltend die Schriftzeichen und Alphabete aller Zeiten und aller Völker des Erdkreises*. Verlag der Kaiserlich-königlichen Hof- und Staatsdruckerei, Wien, 2nd edition, 1880.
- [5] G. Gnoli and ʿA.-A. Saʿīdī Sīrjānī. Entry: Bang. *Encyclopedia Iranica*, III(7):689–691, 1988.
- [6] W. B. Henning. *Zoroaster: Politician or Witch-Doctor*. Oxford University Press, London: Geoffrey Cumberlege, 1951.
- [7] J. Michael Herrmann. The Voynich Manuscript is written in natural language: The Pahlavi hypothesis. *arXiv preprint arXiv:1709.01634*, 2017.
- [8] Francis Johnson. *A dictionary of Persian, Arabic, and English*. W<sup>M</sup>. H. Allen and Co., London, 1852.
- [9] D. N. MacKenzie. *A concise Pahlavi dictionary*. Oxford University Press, London, 1986.
- [10] V. S. Rastorgueva. *A short sketch of the grammar of Persian*. Indiana University, Bloomington, Mouton & Co., The Hague, The Netherlands, 1964. Edited by Herbert H. Paper, translated by Steven P. Hill.
- [11] S. Reddy and K. Knight. What we know about the Voynich manuscript. In *Proc. of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, pages 78–86, Madison, WI, 2011. Omnipress, Inc.
- [12] E. Sherwood and E. Sherwood. The Voynich botanical plants, 2008, accessed on 5/9/2017. [http://edithsherwood.com/voynich\\_botanical\\_plants](http://edithsherwood.com/voynich_botanical_plants).
- [13] F. Steingass. *A comprehensive Persian-English dictionary*. Routledge & Kegan Paul, London, 5th edition, 1963.