



ASSOCIATION
des
RÉSERVISTES
du
CHIFFRE
et de la
SÉCURITÉ
de
L'INFORMATION

Nouvelle série
n° 29 – 2001

Document interne à l'Association
Réservé aux adhérents

6. L'hypothèse de la langue slave dans le manuscrit de VOYNICH

L'énigme du manuscrit de VOYNICH¹ conserve ses secrets depuis sa première apparition survenue en Europe centrale à la fin du XVI^e siècle. Étrangement, l'origine géographique de son apparition a été comme ignorée. Il n'a jamais été formulé d'hypothèses liées aux cultures de l'Europe centrale. Et bien au contraire, les analyses se limitèrent aux cultures de l'Europe de l'ouest avec de temps à autre un regard furtif vers l'hébreu ou le grec. La voie de l'hypothèse slavique est ici empruntée et aboutit à une solution cryptographique de laquelle surgissent quelques bribes de décryptement.

Structures des mots.

Dans le manuscrit de VOYNICH, il a été constaté par CURRIER [3] et par TILTMAN [4] que des « mots » ou des « phrases » diffèrent les uns des autres d'un seul symbole, comme « *sa?* » diffère de « *ca?* » par le symbole « *δ* », bien que cette caractéristique se retrouve dans le langage naturel écrit, on ne la constate pas avec la même proportion.

L'application d'une variante du test de HAMMING sur les mots du manuscrit et sur les mots des langues du haut moyen-âge montre les structures binaires ou *binarité*² des mots de chaque langue et ainsi met en valeur les différences remarquées par CURRIER et TILTMAN.

¹ Cf. Bulletin ARCSI n° 28 – 2000, p. 97 et voir aussi [1].

² La binarité est le résultat de la variante du test de Hamming. Ce résultat exprimé avec des 0 et des 1 indique la structure des mots. La binarité est toujours de la dimension du mot ôté de un. Ainsi un mot de deux lettres aura une binarité égale à « 0 » ou à « 1 ».

Vocable	Voynich	Germain	Flamand	Danois	Bulgare
2	0	1	1	1	1
3	10	10	10	10	10
4	110	110	110	110	100
5	1010	1010	1011	1110	1010
6	10110	11010	? ³	11010	10010
7	101100	?	?	?	?

Tableau 1 : Binarité des mots

On lit dans ce tableau que la binarité d'un mot du manuscrit de deux lettres est différente de la binarité des mots de deux lettres de toutes les autres langues naturelles⁴. Mais aussi que la binarité des mots de trois à cinq lettres du manuscrit est identique à celle de la langue germanique.

Imbrications des mots.

Dans le langage les mots sont associés pour former le discours. L'étude de ces associations permet par exemple de déterminer les mots communs à des discours différents [1]. Dans le manuscrit de VOYNICH il existe deux mots incontournables pour la construction des propositions. Ces deux mots «9» et «δ9» sont liés par la lettre «9». On constate que dans la langue germanique les mots correspondants à «9» et «δ9» pourraient être «IN» et «EIN». La langue danoise est quant à elle différente. Elle a pour base trois mots «I», «DE» et «EN» qui ne se substituent pas à «9» et «δ9». Il en est de même pour les mots «DU» et «EN» de la langue flamande. Finalement, l'hypothèse de la langue bulgare est plausible. En effet, ses deux bases «E» et «TE» sont substituables à «9» et «δ9».

Dans le cas où le manuscrit de VOYNICH a été rédigé sans tentative de dissimulation de l'écriture alors il est fort probable que la langue «Germanique» a été employée pour sa rédaction.

³ Le point d'interrogation signifie que la structure est indéterminable par manque de corpus textuel.

⁴ Les langues naturelles étudiées sont le latin, l'espagnol, le portugais, le gallois, le flamand, l'allemand, le danois, le russe et le bulgare. Seuls sont présentés les résultats utiles dans le présent exposé.

Cependant, comment expliquer que les travaux de J.-R. CHILD [2] sur l'hypothèse germanique n'ont pas abouti puisque ce dernier était persuadé « qu'on avait affaire à un langage naturel et non à un cryptogramme » ? Comment expliquer que les mots de deux lettres ont une binarité (Cf. tableau 1) tout autre de celles des langues naturelles et que seule la binarité des mots de trois à cinq lettres est similaire à celle du germain ? Et enfin, comment expliquer que «δ9» soit quarante fois plus présent à la fin des mots et que «EIN» soit quasi équitablement présent au début ou à la fin des mots de la langue germanique ?

En 1921, le professeur NEWBOLD s'était attaché à la brève de phrase du folio 116v « *michi daba multas portas* » laquelle selon lui indiquait que l'accès au texte clair du manuscrit nécessitait de franchir diverses portes [1]. Si effectivement cette énigme est protégée par diverses « portes », que l'on constate jusqu'à présent comme fermées, alors on doit s'attendre à ce que la binarité des mots soit, elle aussi, protégée de la même sorte.

Hypothèse d'une langue slave.

L'analyse du manuscrit a toujours été conduite dans la perspective de l'existence d'un langage sous-jacent d'Europe de l'ouest ou méditerranéen du type grec ou hébreu voire basque, mais toutes les hypothèses se sont révélées infructueuses. Cependant, l'imbrication de mots a révélé que l'emploi d'une langue bulgare est une solution tout aussi plausible et peut-être est-ce là le signe qu'une nouvelle voie, celle de l'Europe centrale – pays d'origine du manuscrit – doit être explorée.

L'hypothèse qu'une langue d'Europe centrale a été employée n'a jamais été soulevée bien que les symboles du manuscrit puissent pour beaucoup dans l'alphabet grec ; lui-même à la source du cyrillique que l'on retrouve par exemple dans un manuscrit appelé « *Радзивиловская летопись* »⁵ dont l'écriture ressemble à celle utilisée dans le manuscrit de VOYNICH.

⁵ Le manuscrit « *radziviloskaia* » est écrit en langue slave et il raconte la formation de la Russie de Kiev. On le nomme aussi « le manuscrit de Koenigsberg ». Il a été écrit à la fin du XV^e siècle. Une des pièces intéressantes du manuscrit concerne l'écriture cursive des notes placées au début de l'ouvrage.

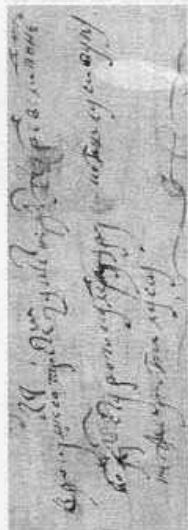


Fig. 1 - Extrait de « Радзивилоская летопись » (Époque moyen-âge).

Bien que l'écriture de ce folio soit difficile à lire, on retrouve dans ce manuscrit slave les lettres « o o z g s ff » du manuscrit de VOYNICH. Des lettres sont plus grandes ou diacrisées et indiquent une sténographie que l'on retrouve communément sur les icônes orthodoxes. Peut-on faire un rapprochement entre ces sténographies et les symboles composés « 2 ff ff » de l'alphabet de Friedman ? Il est à vrai dire trop tôt pour y répondre.

Solution cryptologique.

Les similitudes typographiques entre l'écriture de VOYNICH et le cyrillique ainsi que l'analyse de la structure des mots de VOYNICH ont conduit à la découverte d'une solution cryptographique en faveur de l'hypothèse d'une langue d'Europe centrale de type slave⁶. Cette solution se décompose en trois étapes très classiques en cryptographie : la permutation, l'inversion et la substitution alphabétique.

INVERSION.

Dans le premier tableau on a constaté que la caractéristique des mots de deux lettres du manuscrit est de posséder une binarité inverse de celle des autres langues. Cette apparente symétrie rappelle le problème rencontré par J.-F. CHAMPOLLION sur la détermination du sens de la lecture des hiéroglyphes égyptiens. Quand il remarqua que l'orientation des animaux était liée au sens de la lecture, « tout » lui devint plus clair. On ne

⁶ On étend effectivement la dénomination « langue bulgare » à une « langue slave », le bulgare étant une langue slave.

peut prétendre avoir la même révélation mais il est possible que cette binarité inverse soit le reflet d'une écriture inversée. Il est en effet assuré que la binarité des mots de VOYNICH et de la langue bulgare sont différentes. Pourtant, une simple inversion de la binarité des mots permet de passer de la binarité des mots bulgares à la binarité des mots de VOYNICH.

Vocable	Voynich	Inversion	Bulgare
2	0	0	1
3	10	01	10
4	110	011	100
5	1010	0101	1010
6	10110	01101	10010
7	101100	001101	?

Tableau 2 : Inversion des lettres des mots

Dans ce tableau, l'inversion binaire transforme⁷ chaque « 0 » en « 1 » et chaque « 1 » en « 0 ». Puis, il suffit de lire de droite à gauche la binarité inversée des mots bulgares pour obtenir la binarité des mots de VOYNICH.

PERMUTATIONS.

La binarité est une des caractéristiques du mot. Les différentes positions de substitution sur un mot sont toutes aussi indicatives. En effet, plusieurs mots issus de langages différents et de binarités identiques peuvent avoir des positions différemment utilisées pour la substitution de lettres.

Vocable	Voynich	Bulgare	Binarité identique dans la langue russe
2	2 1	1 2	
3	1 3 2	1 3 2	
4	1 2 4 3	1 4 3 2	
5	1 4 2 5 3	1 5 3 4 2	
6	1 3 2 4 6 5	1 6 5 3 4 2	
7	1 5 2 3 7 6 4	?	

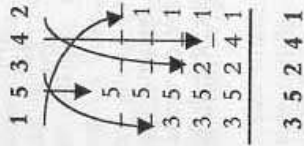
Tableau 3 : priorité de substitutions dans les mots bulgares, russes et de Voynich

⁷ Par exemple, un vocable bulgare de cinq lettres voit sa binarité « 100 » inversée en « 011 ». La lecture de droite à gauche de « 011 » aboutit au vocable de Voynich « 110 ».

On lit dans ce tableau qu'un mot de VOYNICH de cinq lettres a les priorités de substitution « 14253 ». Cela signifie que la première position d'un mot de cinq lettres est la plus utilisée et que la quatrième position de ce mot est la moins utilisée des cinq. Dans la langue bulgare et dans la langue russe un mot de cinq lettres aura les priorités de substitution « 15342 ». Le passage de la combinaison « inversée » d'un mot du manuscrit à la combinaison d'un mot bulgare s'explique grâce à quatre règles de permutation.

Règles de la transposition.

- 1) La position « 2 » reste à la position « 2 » du vocable.
- 2) La position « 1 » va en dernière position « n ».
- 3) La position « 3 » va en position « 1 » du vocable.
- 4) De la droite vers la gauche, la position « 3 < p < n » va à la position « p-2 » s'il y a une place de libre sinon elle va à la dernière place restante.



Le mot bulgare « 15342 » est ici transposé en « 35241 ». Ce résultat est l'inverse du mot de VOYNICH. L'opération d'inversion vue précédemment permet de retrouver la bonne combinaison « 14253 ». Pour autant, il existe des cas où ni la transposition et ni l'inversion ne sont employées pour cacher les mots.

Constellation de Pégase.

Le folio astrologique représentant le signe zodiacal des Poissons contient dans son cœur deux poissons dont un filament s'étire jusqu'à une étoile nommée « $\text{ff}\cdot\text{2}\cdot\text{as}$ ». L'expression de la direction symbolisée par le filament se dirigeant vers une étoile laisse penser que la lettre est « o » un prépositionnel locatif employé pour indiquer sur qui ou quoi porte le propos. L'étiquette « $\text{ff}\cdot\text{2}\cdot\text{as}$ » serait alors composée de deux mots « o » et « $\text{ff}\cdot\text{2}\cdot\text{as}$ ». On constate dans ce mot que les deux lettres très caractéristiques du cyrillique sont « ff » et « 2 ». Aussi, et très

simplement, la recherche d'un mot commençant par la lettre « ff », composé de cinq lettres et relatif à un objet stellaire, aboutit au nom de « nerac ». La surprise fut d'autant plus grande que la constellation de Pégase se trouve dans le zodiaque des Poissons en plein mois de mars.



Fig. 2 - Extrait du folio du signe zodiacal des poissons.

Les lettres cyrilliques associées à l'étiquette « $\text{ff}\cdot\text{2}\cdot\text{as}$ » sont alors « o nerac ».

Voynich	o	ff	2	a	s	
Cyrillique	o	n	e	r	a	c

Il est constaté qu'en ce sens tous les mots du manuscrit ne sont pas cryptés et qu'apparemment il est possible d'entrevoir du texte clair.

Un château.

Si une partie des mots peut apparaître dans le bon ordre, une partie est désordonnée par inversion et par permutation. On

le constate dans un des folios représentant une carte géographique. Sur cette carte est dessiné un château. Un cercle de texte l'entoure. Un mot du texte est « ςοδσλωδ ». Lorsqu'on l'inverse on obtient « ωδσλσς » et lorsqu'on le transpose il devient « ςσλωδσδ ».

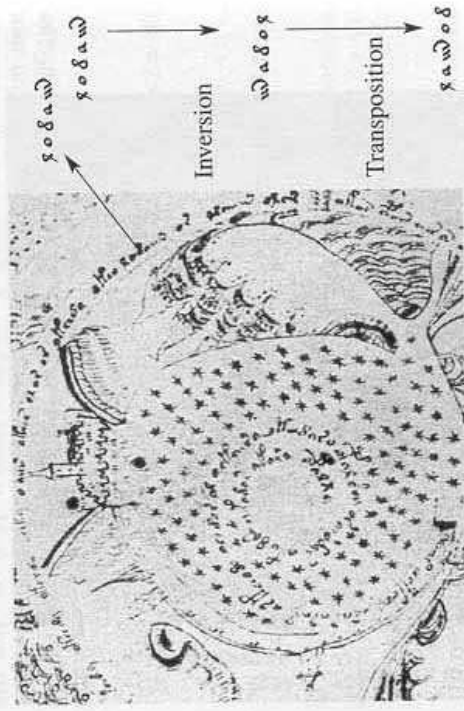


Fig. 3 - Vue sur le château du folio 85-86.

Il n'existe pas de mot commençant par « ςσ » dans le cas où « ς » représenterait la lettre « e ». Aussi, on suppose que les substitutions sont utilisées comme dans le cas de « ο » « ϣε2σς ». Seules les lettres, « ο », « ϣ », « α », n'ont pas été altérées et l'on peut faire l'hypothèse que dans « ςσλωδσδ » les lettres « α » et « ο » ne sont pas représentatives d'autres lettres. La lettre « ς » est la plus proche de la lettre « e » mais dans ce cas il n'existe pas de mot commençant par « ςσ ». Il est probable que la lettre « ς » est ici le résultat d'une substitution comme dans « ϣε2σς ». On exprime le même doute quant à la valeur de la lettre « δ » que l'on ne trouve pas dans l'alphabet cyrillique. La lettre comprise entre « α » et « ο » est « ωδ » laquelle figure dans l'alphabet cyrillique et est équivalente à la lettre « m » de la langue française. Ainsi, on recherche un mot de cinq lettres

dont le cœur est occupé par « αωδσ ». Le nombre de mot est restreint à un seul dans la langue russe⁸ dont le sens dépend de l'accentuation faite sur le « α » ou sur le « ο » ; le mot « замок » pouvant désigner un « Château » ou une « Serrure ».

Voynich ς α ω δ
Cyrillique з а м о б к

Dans la langue bulgare le mot « замк » a la même origine que « замок » et signifie lui aussi « Château ».

Fragment de phrase.

Les mots de trois lettres ne sont pas affectés par l'inversion et la transposition⁹. Cette caractéristique permet d'étudier ces mots directement.

Dans le folio 79v une femme allongée sur le dos reçoit une substance sur le visage. Dans le premier paragraphe de ce folio se trouve un mot de trois lettres aisément distinguable « ςσς ». Il existe en bulgare un verbe « ςσ » dont la troisième conjugaison au singulier est « ςσ ». Ce verbe signifie en français « Gaver » et est ici conjugué en « Il ou elle Gave ».



Fig. 4 - Extrait du folio 79v, vue sur le premier paragraphe.

« замок », « самок », « самок », « самок ». Remarquons que le mot russe « замок » existe aussi en bulgare. La différence repose sur la prononciation : Le « o » devient « b » et donne le mot bulgare « замок » qui signifie « Château ». Un autre mot bulgare « самок » est théoriquement possible, cependant la signification de ce génitif pluriel « des femelles » est « hors contexte » par rapport au dessin.

⁹ Bien qu'ils puissent l'être des substitutions.

On étend notre recherche à un mot voisin afin de vérifier si notre hypothèse se confirme. À la gauche de «*2os*» se trouve un autre mot «*4o ff a w*». L'application des règles de permutations sur ce mot fournit le mot qui inversé devient «*4 a w o ff*». Est-ce à nouveau le nom de «*château*» qui apparaît dans ce paragraphe ? On peut le supposer. Cependant, l'autre mot bulgare plausible est «*camok*». Il se traduit par un génitif pluriel «*des femelles*». Ainsi, la bribe «*4o ff a w 2os*» est transcrite en «*des femelles gave*».

Voynich	4	a	w	o	ff	2	o	s
Cyrillique	C	a	M	O	K	r	O	e

Il reste à déterminer le sujet et le complément de la phrase. Actuellement, on peut comprendre que quelque chose en provenance «*des femelles*», «*gave*» quelque chose ou quelqu'un.

Conclusion.

La structure des mots révèle que la langue de VOYNICH pourrait être une langue slave dissimulée derrière trois portes cryptographiques que sont la transposition, l'inversion et la substitution de lettres, mais parfois, les mots clairs sont simplement cachés par la substitution d'une partie de leurs lettres : il existe donc deux cryptographies différentes. Cette nouvelle voie montre une étoile nommée «*Pégase*». Plus loin, surgit un «*Château*». Et encore plus loin une bribe de phrase apparaît pour dire que quelque chose issu «*des femelles gave*» quelque chose d'autre ; et ces «*petites choses*» restent à découvrir pour la confirmation des décryptements effectués. Déjà, la découverte de ces quatre mots permet de commencer la reconstitution de l'alphabet de substitution.

Voynich	a	c	s	δ	2	w	o	ff	4
	s								

Cyrillique	a	e	3	k	r	M	o/b	n	c
------------	---	---	---	---	---	---	-----	---	---

Cet alphabet indique que l'auteur a utilisé un procédé de substitutions avec changement de référence. Le procédé employé

ouvre la possibilité à des lettres claires d'être représentées par la même lettre de VOYNICH ce qui ne permet pas la réciprocity directe entre une lettre cryptée et une lettre claire. De ce fait, on doit s'attendre à découvrir des règles de changement d'emploi des lettres de VOYNICH lesquelles seront probablement liées aux cinq voire six alphabets différents¹⁰ de lettres utilisés dans cette énigme.

Antoine CASANOVA,
THALES-Communications

Références.

- [1], Antoine CASANOVA, thèse de doctorat, université PARIS-8 (France), méthode d'analyse du langage crypté : une contribution à l'étude du manuscrit de Voynich, Paris, 1999.
- [2], James R. CHILD, *The Voynich manuscript revisited, US Government, Confidential house organ*, 16 février 1976. Complété de notes manuscrites de l'auteur.
- [3], Capitaine Prescott H. CURRIER, *Some Important New Statistical Findings, Seminar on 30th November in Washington D.C., 1976*.
- [4], John H. TILTMAN, *Interim report on the Voynich MS : Personal communication to W.-F. FRIEDMAN*, 5 mai 1951.
- [5], George Kingsley ZIPF, *The psycho-biology of language, Hought Mifflin Co, Boston*, 1935.

¹⁰ Le manuscrit de VOYNICH n'a pas été écrit uniformément avec le même alphabet de lettres. Il a été distingué cinq voire six alphabets différents de lettres [1]. De telle sorte qu'on peut penser que la non réciprocity directe entre une lettre cryptée et une lettre claire s'explique par les changements d'utilisation de ces différents alphabets.