

Table 2: Word-initial character frequency

char	red word				green word			
	> 50	> 20	> 10	> 5	> 50	> 20	> 10	> 5
o	0	0	57	144	1199	1264	1264	1264
4o	0	0	11	11	0	0	0	0
o	0	23	372	705	2744	3387	3494	3494
o	0	0	0	6	0	0	0	0
o	54	54	94	124	68	97	97	97
o	0	0	15	44	0	0	19	19
o	0	0	103	145	244	291	291	291
o	0	0	135	262	2319	2543	2600	2600
o	0	53	82	134	0	0	0	0
o	0	25	37	57	0	0	0	0
o	0	0	11	11	0	0	0	0
o	0	0	15	15	0	0	0	0
o	0	30	30	30	0	0	0	0
o	0	0	0	6	0	0	0	0
o	4299	4774	5110	5328	0	0	0	0
o	1245	1500	1752	1939	347	372	372	372
o	0	28	28	28	0	0	0	0
o	0	0	11	11	57	57	57	57
o	108	247	376	501	0	32	32	32
o	0	0	0	0	0	23	23	23
o	421	617	658	700	101	101	101	101
o	0	0	0	52	643	745	765	765
o	0	0	208	396	1498	1911	1911	1911
o	2469	2665	2876	3125	0	0	0	0
o	0	0	273	454	211	549	586	586

2.2 Observation of initial character

I investigated word-initial character of words in each group. In this analysis, Currier-like character definition is employed. Table 2 shows the frequency of word which begins with each character. From this result, it is found that distributions of the initial character for both groups are very different. Table 3 shows the relationship between initial character and a group the word belongs to. Relationship between red group and its initial character is very similar to property of 4o-prefixed word discovered by Landini[4] and Sazonov[5].

2.3 Observation of final character

Next, word-final character frequencies are investigated. The result is shown in table 4. From this result, it is found that word-final characters have almost nothing to do with the word's group. This result is consistent with the fact that the right entropies of words are independent to their group.

3 Reproductivity of red and green words

Next, reproductivity of words in both groups are investigated. Here, I call a Voynichese word is reproductive when both prefixed and non-prefixed form of the word appears in VMS. I investigated the reproductivity of words by following procedure:

1. Pick a word w . If frequency of the word is more than the threshold, continue processing.
2. if the word is prefixed by o or $4o$, remove the prefix. Let the prefix-removed (or original, if it doesn't have a prefix) word be w' .

Table 3: Relationship between group and initial character

group	characters
red	Ɱ Ɱ Ɱ Ɱ
probably red	c ȝ ʳ ʷ ʷ
green	o c c ȝ ʳ c
probably green	Ɱ o ȝ
not sure	c Ɱ Ɱ Ɱ ȝ ʷ ȝ 4

3. Check if w' , ow' and $4ow'$ appears in the original text.

Table 5 shows the reproductivity of frequent words. It is obvious that red words have higher reproductivity, which means that red words can be prefixed by o and $4o$. On the other hand, reproductivity of green words against o and $4o$ is not very high, but most frequent words are still reproductivity.

References

- [1] Rene Zandbergen: *The Voynich Manuscript*, <http://www.voynich.nu/>
- [2] Akinori Ito: *Observation of left and right entropy in Voynich MS*, <http://www.geocities.co.jp/Technopolis/7220/voy/paper021209.pdf>
- [3] Akinori Ito: *Effects of line context and prefix 4o upon contextual deviation of Voynich 'words'* <http://www.geocities.co.jp/Technopolis/7220/voy/paper021217.pdf>
- [4] G. Landini: *On the structure of the labels and the function of the character <q>*, <http://web.bham.ac.uk/G.Landini/evmt/qowords.htm>
- [5] V. Sazonov: *More about o- and qo-*, <http://voynich.naobum.de/qo.htm>

Table 4: Word-final character frequency

char	red word				green word			
	> 50	> 20	> 10	> 5	> 50	> 20	> 10	> 5
◦	0	0	0	10	0	0	0	0
⊙	0	0	79	177	0	67	67	67
⊙	208	416	586	817	637	743	743	743
◦	0	0	12	12	0	24	24	24
◦	1003	1071	1180	1293	1064	1165	1165	1165
◦	704	735	749	787	137	192	192	192
◦	0	0	17	44	0	0	0	0
◦	0	0	11	11	0	0	0	0
◦	0	0	35	50	0	41	41	41
◦	840	893	1004	1156	1699	1946	2021	2021
◦	682	727	794	880	376	376	395	395
◦	0	0	26	39	0	0	20	20
◦	0	0	0	6	0	0	0	0
◦	0	0	11	17	0	0	0	0
◦	110	160	225	344	169	169	187	187
◦	0	0	15	29	0	24	24	24
◦	0	84	84	116	154	176	176	176
◦	1195	1458	1769	2011	1261	1542	1558	1558
◦	0	0	14	23	0	0	0	0
◦	1804	1968	2326	2554	1540	1788	1788	1788
◦	0	113	258	342	301	368	368	368
◦	0	25	108	210	256	381	381	381
◦	0	0	11	11	0	0	0	0
◦	0	0	26	26	0	0	0	0
◦	2050	2366	2914	3263	1837	2370	2462	2462

Table 5: Reproductivity of words

	prefix	red word				green word			
		> 50	> 20	> 10	> 5	> 50	> 20	> 10	> 5
frequency-weighted	without prefix	100.0	99.7	99.3	98.8	100.0	100.0	100.0	100.0
	with ◦	93.0	94.0	87.4	82.2	87.8	79.2	78.4	78.4
	with 4◦	92.3	90.6	83.2	76.8	75.2	66.1	65.5	65.5
not frequency-weighted	without prefix	100.0	99.0	97.7	96.6	100.0	100.0	100.0	100.0
	with ◦	94.2	96.9	70.5	59.2	79.0	58.1	56.2	56.2
	with 4◦	92.3	84.5	60.5	47.1	64.5	41.1	40.9	40.9