THE BEALE CIPHER: A DISSENTING OPINION

by

James J. Gillogly

This paper originally appeared in Cryptologia, April 1980; Volume 4, Number 2.

Other Papers by Gillogly

- Ciphertext-only Cryptanalysis of Enigma
- Cryptograms from the Crypt
- Breaking an 18th Century Shorthand System

Abstract. The Beale Treasure Cipher (B1) has withstood the attacks of several generations of amateur and professional cryptanalysts. This paper reports a statistical anomaly in B1 which suggests that it my be a hoax.

Keywords: Beale Cipher, homophonic cipher, book cipher, treasure cipher.

In 1885 James B. Ward of Virginia published a pamphlet describing a fabulous treasure buried by an explorer named Thomas Jefferson Beale in Bedford County, Virginia, over 60 years earlier. The location, contents, and intended beneficiaries of the treasure were concealed in three separate ciphers. Ward claimed to have broken the second cipher (B2), describing the contents, and found it to be a book cipher based on the Declaration of Independence (DOI). The words of the DOI were numbered consecutively, and each plaintext letter was replaced with the number of a word in the DOI beginning with that letter. The details of the encryption are discussed exhaustively by Dr. Carl Hammer [1], The initials of words in the DOI are given in Table 1.

The first of the Beale Cipher papers (Bl) contains 495 numbers from 1 to 2906 (Table II). This would seem to preclude the DOI, with only 1322 words, from being the key text. This impression is supported if the first few characters of B1 are decrypted with the DOI, yielding SCS?E TFA?G CDOTT .... where ? stands for the plaintext of a ciphertext number greater than 1322. This much gibberish was probably adequate to dissuade early cryptanalysts from pursuing this line.

But it is difficult to bore a computer. I wrote a very simple program which accepts as input a cryptogram in "Beale cipher" and the initial letters of any document, and attempts the decryption. Table III shows the result of applying the DOI to B1. If the DOI is the wrong key, the resultant text should be a random sequence of letters drawn from the distribution of DOI initials.

There are a number of oddities in this "decryption," but the most striking is the sequence ABFDE FGHII JKLMM NOHPP. Note in passing that the first F is encrypted as 195 and that letter 194 of the DOI is a C. Similarly, the last H is 301, and letter 302 of the DOI is an O. Hammer [1] noted 23 examples where the person who encrypted B2 made errors of this type, or about one every 33 letters. But correcting these errors is not

critical to the argument. We will henceforth consider only the 14-letter monotonically increasing string DEFGH IIJKL MMNO.

This is obviously an unlikely occurrence to tied in an assumed random text. To establish just how unlikely, consider the following simple model: assume 26 letters of equal frequency and find the expected number of monotonic runs of each length. For a sequence of length 3, say, the probability that the second letter is equal to or one greater than the first is 2/26 or 1/13, assuming for the sake of neatness that A is the successor of Z. Similarly, the probability that the third letter continues the sequence is 1/13, so that the probability of a sequence at least 3 letters long is 1/(13^2), or about 10^-4. Thus one would expect to find about three 495/(13^2) sequences of at least 3 characters in a random text of 495 equally likely letters.

Continuing the argument, the probability of a sequence of at least 14 letters is 1/(13^13), or about 10^-14. In a random text of 495 characters, the odds against getting a sequence this long would be about 10^12 to 1.

Those figures are only approximate because the frequencies of initial letters in the DOI are assuredly not evenly distributed across the alphabet. The prevalence of T (the, that, etc.) suggests that strings of T's would be considerably more frequent. Hammer [l] shows 19% for T, which would give the sequence TTTTT (which occurs at position 135) an expected value of 0.6 occurrences in a text this long, which is acceptably high. On the other hand, J and K are very much less common. In order to construct the sequence DEFGH IIJKL MMNO, the hypothetical random selection had to choose one of the 10 J's in the DOI, followed by one of the 4 K's. The effects of the unevenness of the distribution tend to offset one another.

How could this kind of sequence occur? Among the possibilities is that it is a random event, and "just happened" in a cryptogram enciphered using another document. This is quite unlikely, as the previous arguments show. Another possibility is that the DOI is in fact the key, but that another level of encryption (e.g. elimination of nulls) must be stripped away. My investigations do not preclude this possibility, although I have been unable to extract any intelligible plaintext from it. Also, Hammer [3] is convinced that the same method was used to encrypt B1 and B2, and B2 did not use a second level of encryption.

My inclination is to a third possibility: that at least the first document, B1, is a hoax. I visualize the encryptor selecting numbers more or less at random, but occasionally growing bored and picking entries from the numbered Declaration of Independence in front of him, in several cases choosing numbers with an alphabetic sequence.

The view of the Beale ciphers as a hoax is supported to some extent by the decrypted message of B2 [2], which ends "Paper number one describes the exact locality of the vault, so that no difficulty will be had in finding it." Hammer has shown [1] that encryption was, for the author of B2, extremely laborious and fraught with error. Why would he waste the effort of encrypting another 87 characters of a message which would be redundant when the first paper, B1, was deciphered? When viewed as a hoax it makes perfect sense: the author wanted to sell the idea that the first document was worth reading.

It is often much more difficult, if not impossible, to prove that a document is meaningless than to extract the sense from a meaningful one. The observations in this paper do not constitute an unequivocal proof that the Beale treasure cipher, Bl, is a hoax, but they do constitute strong evidence that the Declaration of Independence was used to encipher at least the long alphabetic string. This fact should be taken into account in any theory of the authorship and intent of the Beale Ciphers.

REFERENCES

1. Hammer, Carl. 1979. "How Did TJB Encode B2?" *Cryptologia*. 3: 9-15.

2. Innis, P. B. 1964. "The Beale Fortune." *Argosy*. August: 70-71, 82-84.

3. Kahn, David. 1967. *The Codebreakers*. New York: Macmillan. 771-772.

1	WITCOHEIBN	101	LLATPOHTTS	201	AAEHSTMAMD	301	HOTPKOGBIA
11	FOPTDTPBWH	111	TRGAIAMDTJ	211	TSWEASTTRT	311	HORIAUAHID
21	CTWAATAATP	121	PFTCOTGTWA	221	BATFTWTAAB	321	OTEOAATOTS
31	OTETSAESTW	131	FOGBDOTEII	231	WALTOAAUPI	331	TPTLFBSTAC
41	TLONAONGET	141	TROTPTAOTA	241	TSOEADTRTU	341	WHHRHATLTM
51	ADRTTOOMRT	151	IATINGLIFO	251	ADIITRIITD	351	WANFTPGHHF
61	TSDTCWITTT	161	SPAOIPISFA	261	TTOSGATPNG	361	HGTPLOIAPI
71	SWHTTTBSET	171	TTSSMLTETS	271	FTFSSHBTPS	371	USITOTHASB
81	AMACETTAEB	181	AHPIWDTGLE	281	OTCASINTNW	381	OAWSSHHUNT
91	TCWCURTATA	191	SNBCFLATCA	291	CTTATFSOGT	391	ATTHHRTPOL
401	FTAOLDOPUT	501	IOAHRTTPAL	601	OATAAPOTSH	701	AUFPTBAMTT
41l	PWRTRORITL	511	FTETSRITME	611	HEAMONOASH	711	PFAMWTSCOT
421	ARITTAFTTO	521	TATDOTFWAC	621	SOOTHOPAEO	721	IOTSFCOOTW
431	HHCTLBAPUU	531	WHHETPTPOT	631	TSHHKAUITO	731	APOTWFITOU
441	ADFTDOTPRF	541	SFTPOTLFNO	641	PSAWTCOOLH	741	WOCFDUIMCO
451	TSPOFTICWH	551	FRTPOTETMH	651	HATRTMIOAS	751	TBOTBJFTUB
461	MHHDRHRFOW	561	ARTCONAOLH	661	TTCPHHCWOT	761	STBTFPOFAT
471	MFHIOTROTP	571	HOTAOJBRHA	671	SUTAJFTOCA	771	FSOELIANPE
481	HHRFALTASD	581	TLFEJPHHMJ	681	UBOLGHATTA	781	TAAGAEIBSA
491	TCOTBEWTLP	591	DOHWAFTTOT	691	OPLFQLBOAT	791	TRIAOAEAFI
801	FITSARITCF	901	TABWCOCAPS	1001	TOWHPFRITM	1101	AWHCTBTTOO
811	TAOCAOMVLA	911	PITMBAATUI	1011	HTORPHBAOB	1111	CKTDTUWWII
821	AYTFOOGFSO	921	HOACNHHCOF	1021	RIAPWCITMB	1121	OCACTTHBDT
831	OLADTIWPTL	931	CTCOTHSTBA	1031	EAWMDATIUT	1131	TVOJAOCWMT
841	FUIACWHHAG	941	ATCTBTEOTF	1041	BTROAFPNHW	1141	AITNWDOSAH
851	HBDUOOHPAW	951	ABOTFTBTHH	1051	BWIATOBBWH	1151	TAWHTROMEI
861	WAUHHPOSRO	961	HEDIAUAHET	1061	WTFTTTOABT	1161	WIPFWTTROT
871	CBOTADTLOO	971	BOTIOOFTMI	1071	LTEAUJOUWH	1171	USOAIGCAAT
881	PHIATTTLAO	981	SWKROWIAUD	1081	RTOTCOOEAS	1181	TSJOTWFTRO
891	FMTCTWODDA	991	OAAXACIESO	1091	HWHATTNJAM	1191	OIDITNABAO
1201	TGPOTCSPAD	1301	FROTPODMWM				
1211	TTUCAAOROT	1311	PTEOOLOFAO				
1221	BFAISTTAAF	1321	SH				
1231	AATTBCATAP						
1241	CBTATSOGBI						
1251	AOTBTDATAF						
1261	AISTHFPTLW						
1271	CPCAECATDA						
1281	OAATWISMOR						
1291	DAFTSOTDWI						

Table I. Initial letters of words in the Declaration of Independence.

71 194 38 1701 89 76 11 83 1629 48 94 63 132 16 111 95 84 341 975
14 40 64 27 81 139 213 63 90 1120 8 15 3 126 2018 40 74 758 485
604 230 436 664 582 150 251 284 308 231 124 211 486 225 401 370
11 101 305 139 189 17 33 88 208 193 145 1 94 73 416 918 263 28
500 538 356 117 136 219 27 176 130 10 460 25 485 18 436 65 84 200
283 118 320 138 36 416 280 15 71 224 961 44 16 401 39 88 61 304
12 21 24 283 134 92 63 246 486 682 7 219 184 360 780 18 64 463
474 131 160 79 73 440 95 18 64 581 34 69 128 367 460 17 81 12 103
820 62 116 97 103 862 70 60 1317 471 540 208 121 890 346 36 150
59 568 614 13 120 63 219 812 2160 1780 99 35 18 21 136 872 15 28
170 88 4 30 44 112 18 147 436 195 320 37 122 113 6 140 8 120 305
42 58 461 44 106 301 13 408 680 93 86 116 530 82 568 9 102 38 416
89 71 216 728 965 818 2 38 121 195 14 326 148 234 18 55 131 234
361 824 5 81 623 48 961 19 26 33 10 1101 365 92 88 181 275 346
201 206 86 36 219 320 829 840 68 326 19 48 122 85 216 284 919 861
326 985 233 64 68 232 431 960 50 29 81 216 321 603 14 612 81 360

```
36 51 62 194 78 60 200 314 676 112 4 28 18 61 136 247 819 921
1060 464 895 10 6 66 119 38 41 49 602 423 962 302 294 875 78 14
23 111 109 62 31 501 823 216 280 34 24 150 1000 162 286 19 21 17
340 19 242 31 86 234 140 607 115 33 191 67 104 86 52 88 16 80 121
67 95 122 216 548 96 11 201 77 364 218 65 667 890 236 154 211 10
98 34 119 56 216 119 71 218 1164 1496 1817 51 39 210 36 3 19 540
232 22 141 617 84 290 80 46 207 411 150 29 38 46 172 85 194 36
261 543 897 624 18 212 416 127 931 19 4 63 96 12 101 418 16 140
230 460 538 19 27 88 612 1431 90 716 275 74 83 11 426 89 72 84
1300 1706 814 221 132 40 102 34 858 975 1101 84 16 79 23 16 81
122 324 403 912 227 936 447 55 86 34 43 212 107 96 314 264 1065
323 328 601 203 124 95 216 814 2906 654 820 2 301 112 176 213 71
87 96 202 35 10 2 41 17 84 221 736 820 214 11 60 760
```

Table II. The Beale Treasure Cipher (B1).

```
SCS?E TFA?G CDOTT UCWOT WTAAI WDBII DTT?W TTAAB BPLAA ABWCT
LTFIF LKILP EAABP WCHOT OAPPP MORAL ANHAA BBCCA CDDEA OSDSF
HNTFT ATPOC ACBCD DLBER IFEBT HIFOE HUUBT TTTTI HPAOA ASATA
ATTOM TAPOA AAROM PJDRA ??TSB COBDA AACPN RBABF DEFGH IIJKL
MMNOH PPAWT ACMOB LSOES SOAVI SPFTA OTBTF THFOA OGHWT ENALC
AASAA TTARD SLTAW GFESA UWAOL TTAHH TTASO TTEAF AASCS TAIFR
CABTO TLHHD TNHWT STEAI EOAAS TWTTS OITSS TAAOP IWCPC WSOTT
IOIES ITTDA TTPIU FSFRF ABPTC COAIT NATTO STSTF ??ATD ATWTA
TTOCW TOMPA TSOTE CATTO TBSOG CWCDR OLITI BHPWA AE?BT STAFA
EWCI? CBOWL TPOAC TEWTA FOAIT HTTTT OSHRI STEOO ECUSC ?RAIH
RLWST RASNI TPCBF AEFTB
```

Table III. "Decryption" of B1 using the DOI.

Mail Jim Gillogly

Converted to hypertext by Joe Peschel October, 2000.