

Kryptos

by

Kristine Seaman, B.A.

A Thesis

In

Mathematics

Submitted to the Graduate Faculty  
of Texas Tech University in  
Partial Fulfillment of  
the Requirements for the Degree of

Master of Science

Approved

Mara D. Neusel  
Chair of Committee

Chris Monico

Lars W. Christensen

Peggy Gordon Miller  
Dean of the Graduate School

May, 2012

©2012, Kristine Seaman

## ACKNOWLEDGEMENTS

I wish to acknowledge the help and support of Dr. Mara D. Neusel and Dr. Chris Monico. Dr. Neusel came up with the idea of working on Kryptos. We were both sitting in her office discussing a cryptology book that we were reading through together when she pulled out the *New York Times* article about Kryptos. Cryptology has always been an interest of mine and Kryptos seemed like an interesting sculpture, and sure enough it was! Thank you Dr. Neusel for giving me this topic which I have enjoyed studying. I greatly appreciate all of your time, help, and support. Dr. Monico has made available his support in a number of ways. Before Dr. Neusel and I started working on the Kryptos sculpture, Dr. Monico had already begun working through possibilities on how to crack it. Throughout my thesis I have used some of his research including his frequency count data. He has always been willing to help me when needed. I would like to thank him for his countless hours of patience and teaching. I have learned a lot from Dr. Monico and Dr. Neusel and I owe them my sincerest thanks. Further, since the sculpture is located in the CIA headquarters building's courtyard, I would not have been able to provide photographs myself. So, I would like to thank Mr. Jim Gillogly for allowing me to use his photographs throughout my study. I would also like to thank my parents, Jody and David Seaman. They have been very inspirational and by my side through it all. Their constant encouragement and praise has helped me in completing this thesis.

## CONTENTS

Acknowledgements . . . . .	ii
Abstract . . . . .	iv
List of Tables . . . . .	v
List of Figures . . . . .	vi
1. Introduction . . . . .	1
2. Shift Ciphers . . . . .	6
3. Vigenère Ciphers . . . . .	7
3.1 Vigenère Cipher with an Algebraic Portrayal . . . . .	9
3.2 How Can a Vigenère Cipher be Broken? . . . . .	10
3.3 The First Section of Kryptos . . . . .	14
3.4 The Second Section of Kryptos . . . . .	15
4. Transpositions . . . . .	18
4.1 Columnar Transpositions . . . . .	23
4.2 The Third Section of Kryptos . . . . .	24
5. The Fourth Section of Kryptos . . . . .	27
5.1 Clues . . . . .	27
5.2 ABCD . . . . .	28
5.3 Running Key and One Time Pad . . . . .	32
6. Conclusion . . . . .	35
Bibliography . . . . .	36

## ABSTRACT

This thesis is about cryptology and the Kryptos sculpture in Langley, Virginia. Kryptos is designed with a coded message that has not been completely solved for over 20 years. We are going to explain what we know about Kryptos and the processes to solve the first three parts of the sculpture. We will cover the shift cipher, the Vigenère cipher, and the transposition cipher. Then we will discuss several methods that the sculptor, Jim Sanborn, may have used to code his fourth message.

## LIST OF TABLES

1.0.1 Kryptos Sculpture Text Cutout . . . . .	5
3.0.1 Standard Vigenère Tableau . . . . .	8
3.0.2 Kryptos Alphabet Vigenère Tableau . . . . .	9
3.2.3 Index of Coincidence for Possible Key Lengths . . . . .	14
4.0.1 English Letter Frequency . . . . .	19
4.0.2 The top 77 digrams occurring within words in the sample, with percentages (including blanks) . . . . .	21
4.0.3 The top 77 digrams occurring within words in the sample, with percentages (without blanks) . . . . .	21
4.0.4 The top 77 most common trigrams within English words, with percentages (including blanks) . . . . .	22
4.0.5 The top 77 most common trigrams within English words, with percentages (without blanks) . . . . .	22
5.2.1 Index of Coincidence for Possible Key Lengths . . . . .	29
5.2.2 Comparison of Letter Frequencies on the Fourth Section of the Sculpture	31

## LIST OF FIGURES

1.1	Kryptos Sculpture . . . . .	1
1.2	Inside of the Courtyard . . . . .	3
1.3	The Compass Points North . . . . .	3
1.4	Morse Code Message . . . . .	3
1.5	The Sculpture's Cut-Out . . . . .	3
4.1	English Letter Frequency Chart . . . . .	19
4.2	Section 3 Plaintext Written into 28 Columns . . . . .	25
4.3	Section 3 Plaintext New Grid with 21 Columns . . . . .	26
4.4	Section 3 Plaintext Transposed to Ciphertext . . . . .	26

## CHAPTER 1

### INTRODUCTION

Kryptos is a sculpture by Jim Sanborn located in the courtyard at the Central Intelligence Agency headquarters in Langley, Virginia. Kryptos is designed with a coded message that has not been completely solved for over 20 years. Most of the message has been decoded since 1999, but the rest is yet to be revealed. In 1999, a computer scientist from California, Jim Gillogly, announced that he had found the solution to the first three sections using a computer. After Gillogly made his announcement the CIA revealed that one of their analysts, David Stein, had previously solved the same three parts one year earlier using paper and pencil [4].



Figure 1.1. Kryptos Sculpture [10]

Jim Sanborn is an artist with many sculptures focusing on topics such as magnetism, the Coriolis effect, secret messages, and mysteries of atomic reactions. He grew up in Alexandria, Virginia and attended Randolph-Macon College and received a bachelor's degree in paleontology, fine arts, and social anthropology in 1968. He later received his masters degree in sculpture from Pratt Institute [3]. When he was chosen for the project, Ed Scheidt, chairman of the CIA's Cryptographic Center, and known within the agency as the "Wizard of Codes,"



taught Sanborn about basic cryptography [16]. Sanborn then spent two and a half years cutting approximately 1,800 individual letters and question marks in rows onto a giant sheet of copper that was planned to be the main piece of the work.

The name Kryptos comes from the Greek word for *hidden* [15]. The theme of the sculpture is apparently “intelligence gathering.” The sculpture is a large sheet of copper resembling a scroll with an encrypted text punched through the copper attached to a piece of petrified wood [23], measuring 12’x20’x10’ [14]. The sculpture’s ciphertext contains at least four separate messages, three of which have been revealed as mentioned above. However, these four messages are not the four sections we are able to see in the photograph above. One side of the sculpture is a key and the other side is the message. We will go into further details later. As the main sculpture was being installed at the CIA, Sanborn also placed several other pieces around the grounds including several large granite slabs and copper sheets. Several Morse code messages are written on the copper and one of the slabs has a compass rose pointing north on it. The Morse code messages that have been noted include SOS, LUCID MEMORY, T IS YOUR POSITION, SHADOW FORCES, VIRTUALLY INVISIBLE, DIGETAL INTERPRETATU, and RQ (or YR if you look from the other direction) [4]. Other elements of Sanborn’s installation include a landscaped area, a duck pond, and several other unmarked rocks. In 1990 the project was installed at the CIA headquarters with a total cost of \$250,000 [15]. In an interview with Sanborn he remarks, “I assumed the code would be cracked in a fairly short time” [15].



Figure 1.2. Inside of the Courtyard [10]



Figure 1.3. The Compass Points North [10]



Figure 1.4. Morse Code Message [10]



Figure 1.5. The Sculpture's Cut - Out [10]

The sculpture seems to look like four panels; we can compare this to the four quadrants on a coordinate plane. If we refer to a coordinate plane (note that this is also affected by which side of the sculpture we are standing) the second and third quadrant would be the ciphertext and the first and fourth quadrant would be the key (for at least for two parts of the ciphertext). However the key would be read from the other side so the letters would appear backwards when we are standing on this side of the sculpture. Or similarly, if we were standing on the other side of the sculpture and refer to a coordinate plane, the second and fourth quadrant would be the key and the first and third quadrants would be the ciphertext (and the ciphertext would now appear backwards). Also, each section of the ciphertext is encrypted differently. The following is a transcription of the sculpture itself, obtained from [5].

Table 1.0.1. Kryptos Sculpture Text Cutout

EMUFPHZLRFAXYUSDJKZLDRNSHGNFIVJ	ABCDEFGHIJKLMNQRSTUUVWXYZABCD
YQTQUXQBQVYUUVLLTREVJYQTMKYRDMFD	AKRYPTOSABCDEFGHIJKLMNUVWXZKRYP
VFPJUDEEHZWETZYVGWHKKQETGFQJNCE	BRYPTOSABCDEFGHIJKLMNUVWXZKRYPT
GGWHKK?DQMCPFQZDQMMIAGPFXHQRLG	CYPTOSABCDEFGHIJKLMNUVWXZKRYPTO
TIMVMZJANQLVKQEDAGDVFRPJUNGEUNA	DPTOSABCDEFGHIJKLMNUVWXZKRYPTOS
QZGZLECGYUXUEENJTBJLBQCRTBJDFHRR	ETOSABCDEFGHIJKLMNUVWXZKRYPTOSA
YIZETKZEMVDUFKSJHKFWHKUWQLSZFTI	FOSABCDEFGHIJKLMNUVWXZKRYPTOSAB
HHDDDUVH?DWKBFUFPWNTDFIYCUQZERE	GSABCDEFGHIJKLMNUVWXZKRYPTOSABC
EVLDKFEZMOQQJLTTUGSYQPFEUNLAVIDX	HABCDEFGHIJKLMNUVWXZKRYPTOSABCD
FLGGTEZ?FKZBSFDQVGOGIPUFXHHDRKF	IBCD EFGHIJKLMNUVWXZKRYPTOSABCDE
FHQNTGPUAECNUVPDJMQCLQUMUNEDFQ	JCDEFGHIJKLMNUVWXZKRYPTOSABCDEF
ELZZVRRGKFFVOEEXBDMVPNFQXEZLGRE	KDEFGHIJKLMNUVWXZKRYPTOSABCDEF
DNQFMPNZGLFLPMRJQYALMGNUVPDXVKP	LEFGHIJKLMNUVWXZKRYPTOSABCDEF
DQUMEBEDMHDAFMJGZNUPLGEWJLLAETG	MFGHIJKLMNUVWXZKRYPTOSABCDEFGHI
ENDYAHROHNLSRHEOCPTIOIBIDYSHNAIA	NGHIJKLMNUVWXZKRYPTOSABCDEFGHIJL
CHTNREYULDSLSSLNOHSNOSMRWXMNE	OHIJKLMNUVWXZKRYPTOSABCDEFGHIJL
TPRNGATIHNRRARPESLNNELEBLPIIACAE	PIJKLMNUVWXZKRYPTOSABCDEFGHIJLM
WMTWN DITEENRAHCTENEUDRETNAEOE	QJKLMNUVWXZKRYPTOSABCDEFGHIJLMN
TFOLSEDTIWENHAEIOYTEYQHEENCTAYCR	RLMNQUVWXZKRYPTOSABCDEFGHIJLMNQ
EIFTBRSPAMHHEWENATAMATEGYEERLB	SMNQUVWXZKRYPTOSABCDEFGHIJLMNQU
TEEFOASF IOTUETUAEOTOARMAEERTNRTI	TNQUVWXZKRYPTOSABCDEFGHIJLMNQUV
BSEDDNIAAHTTMSTEWPIEROAGRIEWFEB	UQUVWXZKRYPTOSABCDEFGHIJLMNQUVW
AECTDDHILCEIHSITEGOEAO SDDRYDLORIT	VUVWXZKRYPTOSABCDEFGHIJLMNQUVWX
RKLMLEHAGTDHARDPNEOHMGFMFEUHE	WVWXZKRYPTOSABCDEFGHIJLMNQUVWXZ
ECDMRIPFEIMEHNLSSTTRTVDOHW?OBKR	XWXZKRYPTOSABCDEFGHIJLMNQUVWXZK
UOXOGHULBSOLIFBBWFLRVQQRNGKSSO	YXZKRYPTOSABCDEFGHIJLMNQUVWXZKR
TWTQSJSSEKZZWATJKLUDIAWINFBNYP	ZZKRYPTOSABCDEFGHIJLMNQUVWXZKRY
VTTMZFPKWGDKZXTJCDIGKUHUAUEKCAR	ABCDEFGHIJKLMNQRSTUUVWXYZABCD

## CHAPTER 2

### SHIFT CIPHERS

We begin with a message, also known as our *plaintext*.

*this is kristine's thesis*

To encrypt our message we can move (or *shift*) each letter, say twice, so our encrypted text (*ciphertext*) will be,

VJKU KU MTKUVKPGU VJGUKU

So,  $t \mapsto V$ ,  $h \mapsto J$ ,  $i \mapsto K$ , and so on. If  $z$  would have been in our plaintext, we would shift it around to the beginning of the alphabet, to **B**. Notice that we changed the font between the plaintext and the ciphertext, this is just to reduce the confusion between the texts.

The *shift cipher* is sometimes also known as the *Caesar Cipher* because Caesar used this method to encrypt his secrets. However, Caesar always shifted three places forward [8, p.2]. Once the method of encryption is recognized, we are able to think of the shift cipher as a guessing game. We must guess the correct number of times to shift a message in order to reveal the correct plaintext. If a ciphertext is long in length, it is not necessary to decrypt the entire message with each shift-length choice. We can normally tell by shifting the first 5 letters if the number we chose to shift was correct. This is because the decrypted message should look like a message read in English. So, if the plaintext appears to be gibberish then it was probably not the correct number of shifts.

If we shift the message say 5 letters forward in the alphabet, to decrypt our message we shift 5 letters backwards. Since knowing the encryption key means we know the decryption key, the shift cipher is known as a *symmetric* cipher. Also, since when using a shift cipher we always encrypt a letter in the same way throughout the message, it is known to be *monoalphabetic* [8, p. 4].

### CHAPTER 3

#### VIGENÉRE CIPHERS

The *Vigenère cipher* can be thought of as a more complex shift cipher. The Vigenère cipher is *symmetric*, meaning that the key to encode the message is the same as the key to decode it, as described in the previous chapter. The Vigenère cipher is also *polyalphabetic*, meaning that a letter in the plaintext may not be encrypted the same way each time that it appears.

Suppose we have our plaintext *this is an example* and we want to encrypt our message with a Vigenère cipher. First we choose our key, say *KRISTINE*. Now we will shift *t*, *K* times. Since K is the 10th letter of the alphabet (we start with A=0, B=1,...) we shift *t* 10 letters, this will result with our first ciphertext letter of *D*. We will now shift *h*, *R* times. Since R is the 17th letter of the alphabet, we shift *h* 17 letters, resulting in a *Y*. We can now repeat this process until we have completely encrypted the entire message. The completed encryption is shown below.

<i>plaintext:</i>	<i>t</i>	<i>h</i>	<i>i</i>	<i>s</i>	<i>i</i>	<i>s</i>	<i>a</i>	<i>n</i>	<i>e</i>	<i>x</i>	<i>a</i>	<i>m</i>	<i>p</i>	<i>l</i>	<i>e</i>
<i>KEY:</i>	<i>K</i>	<i>R</i>	<i>I</i>	<i>S</i>	<i>T</i>	<i>I</i>	<i>N</i>	<i>E</i>	<i>K</i>	<i>R</i>	<i>I</i>	<i>S</i>	<i>T</i>	<i>I</i>	<i>N</i>
<b>CIPHERTEXT:</b>	<b>D</b>	<b>Y</b>	<b>Q</b>	<b>K</b>	<b>B</b>	<b>A</b>	<b>N</b>	<b>R</b>	<b>O</b>	<b>O</b>	<b>I</b>	<b>E</b>	<b>I</b>	<b>T</b>	<b>R</b>

Therefore we have **DYQKBANROOIEITR** as our Vigenère-ciphered message.

A cryptographer may use a *Vigenère tableau* to encrypt their text with a Vigenère cipher. This can make it simpler to complete the encryption, since the tableau reduces the counting factor in the shifting of the letters. Below is what a Vigenère tableau looks like with the standard English alphabet.

Table 3.0.1. Standard Vigenère Tableau

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

If we want to encrypt the text *attack at dawn* with key *CRYPTOLOGY* with the above standard English tableau we would find the letter *a* along the top row and match it with the letter *C* on the left side and meet them where they intersect on the tableau, giving a *C*. Next, we would take the letter *t* along the top row and match it with the letter *R* on the left hand side and we will get a *K*. We can repeat this process until our original text is encrypted.

One side of the Kryptos sculpture is an altered version of the standard Vigenère tableau, see the right hand side of Table 1.0.1. Sanborn started the alphabet with *KRYPTOS* and then listed the remaining letters of the alphabet afterwards. We will refer to this alphabet as the *Kryptos alphabet*. Also notice how there are four extra columns at the end of the alphabet. We will further discuss this when we get to the fourth section of Kryptos. Below we can see the tableau on the sculpture in a clearer chart form. (Note: the extra L in the last column appears on sculpture.)

Table 3.0.2. Kryptos Alphabet Vigenère Tableau

	KRYPTOSABCDEFGHIJLMNQUVWXZKRYPT
K	KRYPTOSABCDEFGHIJLMNQUVWXZKRYPT
R	RYPTOSABCDEFGHIJLMNQUVWXZKRYPT
Y	YPTOSABCDEFGHIJLMNQUVWXZKRYPTO
P	PTOSABCDEFGHIJLMNQUVWXZKRYPTOS
T	TOSABCDEFGHIJLMNQUVWXZKRYPTOSA
O	OSABCDEFGHIJLMNQUVWXZKRYPTOSAB
S	SABCDEFGHIJLMNQUVWXZKRYPTOSABC
A	ABCDEFGHIJLMNQUVWXZKRYPTOSABCD
B	BCDEFGHIJLMNQUVWXZKRYPTOSABCDE
C	CDEFGHIJLMNQUVWXZKRYPTOSABCDEF
D	DEFGHIJLMNQUVWXZKRYPTOSABCDEFG
E	EFGHIJLMNQUVWXZKRYPTOSABCDEFGH
F	FGHIJLMNQUVWXZKRYPTOSABCDEFGHI
G	GHIJLMNQUVWXZKRYPTOSABCDEFGHJL
H	HIJLMNQUVWXZKRYPTOSABCDEFGHJL
I	IJLMNQUVWXZKRYPTOSABCDEFGHJLM
J	JLMNQUVWXZKRYPTOSABCDEFGHJLMN
L	LMNQUVWXZKRYPTOSABCDEFGHJLMNQ
M	MNQUVWXZKRYPTOSABCDEFGHJLMNQ
N	NQUVWXZKRYPTOSABCDEFGHJLMNQ
Q	QUVWXZKRYPTOSABCDEFGHJLMNQ
U	UVWXZKRYPTOSABCDEFGHJLMNQ
V	VWXZKRYPTOSABCDEFGHJLMNQ
W	WXZKRYPTOSABCDEFGHJLMNQ
X	XZKRYPTOSABCDEFGHJLMNQ
Z	ZKRYPTOSABCDEFGHJLMNQ

## 3.1 Vigenère Cipher with an Algebraic Portrayal

Identify the letters A thru Z with the numbers 0 thru 25. To encrypt or decrypt a message we will add (to encrypt) or subtract (to decrypt) modulo 26. That is,

$$\begin{aligned} \text{Plaintext Letter} &\longmapsto \text{Encryption(Plaintext Letter)} \\ &= \text{Plaintext letter} + \text{Key} \pmod{26}. \end{aligned}$$

For example, the letter  $\mathbf{f}$  (5th letter of the alphabet) using the key  $\mathbf{K}$  (10th letter of the alphabet) can be encrypted,

$$5 \longmapsto (5 + 10) \pmod{26} \equiv 15 \quad \text{id est,} \quad \mathbf{f} \longmapsto (\mathbf{f} + \mathbf{K}) \pmod{26} \equiv \mathbf{P}.$$

Likewise, for the decryption of  $\mathbf{P}$  (15th letter of the alphabet) using the key  $\mathbf{K}$



(10th letter of the alphabet) we have,

$$15 \mapsto (15 - 10) \pmod{26} \equiv 5 \quad \text{id est,} \quad \mathbf{P} \mapsto (\mathbf{P} - \mathbf{K}) \pmod{26} \equiv \mathbf{f}.$$

### 3.2 How Can a Vigenère Cipher be Broken?

A weakness of the Vigenère cipher is that it is periodic. Periodicity causes the same letter in the plaintext to occasionally be encrypted the same way twice with a sufficiently long plaintext. A common method for breaking a Vigenère cipher involves first determining the length of the keyword that was used to encipher the message. We can do this by looking for the *index of coincidence*.

The index of coincidence of a given text, abbreviated IC, is the probability that two randomly selected letters will be identical. It can be determined by the formula

$$\sum_{i=1}^{26} \frac{\binom{n_i}{2}}{\binom{t}{2}} = \sum_{i=1}^{26} \frac{n_i(n_i - 1)}{t(t - 1)},$$

where  $t$  is the length of the text and  $n_i$  is the number of occurrences of the  $i^{\text{th}}$  letter [18]. This technique was invented by William F. Friedman in 1922 when he published *Index of Coincidence and its Applications in Cryptology*.

**Example 3.1.** [11, p. 337] *If a message was encrypted with a uniformly random key (encrypted with a one time pad), we can imagine a hat filled with the 26 letters of the alphabet. The chance of pulling out an A is  $1/26$ . If we had two such hats, the probability of pulling out two A's simultaneously is  $\frac{1}{26} \times \frac{1}{26}$ . Further, the chance of a pair of letters (any pair) is the sum of all the probabilities,  $(\frac{1}{26} \times \frac{1}{26}) + \dots + (\frac{1}{26} \times \frac{1}{26})$ , repeated 26 times. This equals .0385.*

A message generated as the one from Example 3.1 is known as *random text* because the letters obtained would be randomly drawn from the hat (or for non example purposes, randomly drawn from a computer).

**Example 3.2.** [11, p. 337] *Imagine a hat filled with 1000 letters of English in the proportion in which they are used in normal English text. This means that the chance of pulling out a specific letter, say A, would be about  $\frac{80}{1000}$ . If we had two such hats, the probability of pulling out two A's simultaneously is  $\frac{80}{1000} \times \frac{80}{1000}$ . Further, the*

chance of a pair of letters (any pair) is the sum of all the probabilities,  
 $(\frac{80}{1000} \times \frac{80}{1000}) + (\frac{10}{1000} \times \frac{10}{1000}) + \cdots + (\frac{2}{1000} \times \frac{2}{1000})$ , for all 26 letters of the alphabet.  
 This equals approximately .0667.

These two plaintext hats can be replaced by two strings of English plaintext. If they are placed one on top of the other, there will be an approximate .0667 probability that two identical letters will have vertical alignment, just as if two identical letters were drawn from a hat.

**Example 3.3.** [11, p. 378]

text A	wheninthe	courseof	humane	events	itbecomes	necessary	for	o	*
text B	fourscore	andseven	years	sago	ourfathers	brought	forth	upo	*
	*	**		*	*				
text A	nenation	to	dissolve	thepolitical	bandsthat	have	connect		
text B	nthis	continent	anewnation	conceived	in	liberty	and	dedic	

There are seven coincidences in 100 pairs (the \* symbol marks the coincidences above); this is precisely what we predicted (6.67%) since 7% of the letters vertically aligned.

This theory also works for enciphered texts. If we superimpose two identically monoalphabetically enciphered English texts we will find an approximately .0667 index of coincidence. This is because with a monoalphabetic substitution, two occurrences of the same plaintext letter result in two occurrences of the same ciphertext letter. Thus, the coincidences will occur in the ciphertext at the same locations they occur in the plaintext. Further, two polyalphabetical enciphered texts with the same key superimposed so that two occurrences of that key are in synchronization with one another will also reveal approximately .0667 index of coincidence [11, p. 378]. This is because the two letters of each vertical pair have the same key letter; whenever a coincidence occurs in the plaintext, the letters of the pair will be enciphered with the same key, resulting as a coincidence in the ciphertext. Thus, the total number of coincidences in the positions of the ciphertext will be the same number of corresponding coincidences in the plaintext. Therefore we can take a single ciphertext and arrange the text into multiple rows. If the average indices of coincidence in each column throughout the ciphertext is approximately .0667, it is likely that we have found the key length of a Vigenère cipher. We can now apply this to the Kryptos sculpture.

**Example 3.4.** *Arrange the ciphertext of the first two rows on the sculpture into 6 columns:*

<i>E</i>	<i>M</i>	<i>U</i>	<i>F</i>	<i>P</i>	<i>H</i>
<i>Z</i>	<i>L</i>	<i>R</i>	<i>F</i>	<i>A</i>	<i>X</i>
<i>Y</i>	<i>U</i>	<i>S</i>	<i>D</i>	<i>J</i>	<i>K</i>
<i>Z</i>	<i>L</i>	<i>D</i>	<i>K</i>	<i>R</i>	<i>N</i>
<i>S</i>	<i>H</i>	<i>G</i>	<i>N</i>	<i>F</i>	<i>I</i>
<i>V</i>	<i>J</i>	<i>Y</i>	<i>Q</i>	<i>T</i>	<i>Q</i>
<i>U</i>	<i>X</i>	<i>Q</i>	<i>B</i>	<i>Q</i>	<i>V</i>
<i>Y</i>	<i>U</i>	<i>V</i>	<i>L</i>	<i>L</i>	<i>T</i>
<i>R</i>	<i>E</i>	<i>V</i>	<i>J</i>	<i>Y</i>	<i>Q</i>
<i>T</i>	<i>M</i>	<i>K</i>	<i>Y</i>	<i>R</i>	<i>D</i>
<i>M</i>	<i>F</i>	<i>D</i>			

*To find the index of coincidence we find the probability of vertical coincidence in each column and then average all of the columns' probabilities. Above we can see that we have:*

<i>Column 1: 2 Z's, 2Y's</i>	<i>Column 4: 2 F's</i>
<i>Column 2: 2 M's, 2 L's, 2 U's</i>	<i>Column 5: 2 R's</i>
<i>Column 3: 2 D's, 2 V's</i>	<i>Column 6: 2 Q's</i>

*Using this information we find the index of coincidence:*

$$\frac{\frac{\binom{2}{2} + \binom{2}{2}}{\binom{11}{2}} + \frac{\binom{2}{2} + \binom{2}{2} + \binom{2}{2}}{\binom{11}{2}} + \frac{\binom{2}{2} + \binom{2}{2}}{\binom{11}{2}} + \frac{\binom{2}{2}}{\binom{10}{2}} + \frac{\binom{2}{2}}{\binom{10}{2}} + \frac{\binom{2}{2}}{\binom{10}{2}}}{6} \approx 0.0323$$

Now lets try 10 columns:

<i>E</i>	<i>M</i>	<i>U</i>	<i>F</i>	<i>P</i>	<i>H</i>	<i>Z</i>	<i>L</i>	<i>R</i>	<i>F</i>
<i>A</i>	<i>X</i>	<i>Y</i>	<i>U</i>	<i>S</i>	<i>D</i>	<i>J</i>	<i>K</i>	<i>Z</i>	<i>L</i>
<i>D</i>	<i>K</i>	<i>R</i>	<i>N</i>	<i>S</i>	<i>H</i>	<i>G</i>	<i>N</i>	<i>F</i>	<i>I</i>
<i>V</i>	<i>J</i>	<i>Y</i>	<i>Q</i>	<i>T</i>	<i>Q</i>	<i>U</i>	<i>X</i>	<i>Q</i>	<i>B</i>
<i>Q</i>	<i>V</i>	<i>Y</i>	<i>U</i>	<i>V</i>	<i>L</i>	<i>L</i>	<i>T</i>	<i>R</i>	<i>E</i>
<i>V</i>	<i>J</i>	<i>Y</i>	<i>Q</i>	<i>T</i>	<i>M</i>	<i>K</i>	<i>Y</i>	<i>R</i>	<i>D</i>
<i>M</i>	<i>F</i>	<i>D</i>							

Above we can see that we have:

Column 1: 2 <i>V</i> 's	Column 6: 2 <i>H</i> 's
Column 2: 2 <i>J</i> 's	Column 7: no coincidences
Column 3: 4 <i>Y</i> 's	Column 8: no coincidences
Column 4: 2 <i>U</i> 's, 2 <i>Q</i> 's	Column 9: 3 <i>R</i> 's
Column 5: 2 <i>S</i> 's, 2 <i>T</i> 's	Column 10: no coincidences

Using this information we find the index of coincidence:

$$\frac{\frac{\binom{2}{2}}{\binom{7}{2}} + \frac{\binom{2}{2}}{\binom{7}{2}} + \frac{\binom{4}{2}}{\binom{7}{2}} + \frac{\binom{2}{2} + \binom{2}{2}}{\binom{6}{2}} + \frac{\binom{2}{2} + \binom{2}{2}}{\binom{6}{2}} + \frac{\binom{2}{2}}{\binom{6}{2}} + \frac{\binom{3}{2}}{\binom{6}{2}}}{10} \approx 0.0914$$

Chance alone will produce a probability of .0385, and monoalphabetic and polyalphabetic cryptograms of English text produce a probability of .0667 [11, p. 378, 379]. Notice the 10 columns of text has an IC closer to .0667. Therefore, it is more probable that the key length is a multiple of 10 rather than 6 when comparing the two.

By calculating all likely key lengths' index of coincidence we can determine a probable key length to decrypt a Vigenère Cipher. In Table 3.2.3 below we determined the index of coincidence for several different key lengths for the first two rows on the Kryptos sculpture. The index of coincidence that is closer to the probability of .0667 has a higher chance of being the correct key length if it was encrypted with a Vigenère cipher.

Table 3.2.3. Index of Coincidence for Possible Key Lengths

key length	1	2	3	4	5	6
IC	0.0379	0.0394	0.0365	0.0345	0.0790	0.0323
key length	7	8	9	10	11	12
IC	0.0357	0.0089	0.0317	0.0914	0.0273	0.0250

Notice how all of the key lengths besides 5 and 10 have an index of coincidence close to .0385, the approximate value of a random text. At 5 and 10 we see an index of coincidence close to .0667, the approximate value of an English text. This provides us with the assertion that our keyword has a probable length of five or ten.

### 3.3 The First Section of Kryptos

Sanborn used a Vigenère cipher with *PALIMPSEST* as the keyword for the first section of Kryptos (abbreviated K1). One meaning of palimpsest, according to Webster’s New Collegiate Dictionary [9, p. 818], is a writing material (as a parchment or tablet) used one or more times after earlier writing has been erased.

<i>P</i>	<i>A</i>	<i>L</i>	<i>I</i>	<i>M</i>	<i>P</i>	<i>S</i>	<i>E</i>	<i>S</i>	<i>T</i>		<i>b</i>	<i>e</i>	<i>t</i>	<i>w</i>	<i>e</i>	<i>e</i>	<i>n</i>	<i>s</i>	<i>u</i>	<i>b</i>
<i>E</i>	<i>M</i>	<i>U</i>	<i>F</i>	<i>P</i>	<i>H</i>	<i>Z</i>	<i>L</i>	<i>R</i>	<i>F</i>		<i>t</i>	<i>l</i>	<i>e</i>	<i>s</i>	<i>h</i>	<i>a</i>	<i>d</i>	<i>i</i>	<i>n</i>	<i>g</i>
<i>A</i>	<i>X</i>	<i>Y</i>	<i>U</i>	<i>S</i>	<i>D</i>	<i>J</i>	<i>K</i>	<i>Z</i>	<i>L</i>		<i>a</i>	<i>n</i>	<i>d</i>	<i>t</i>	<i>h</i>	<i>e</i>	<i>a</i>	<i>b</i>	<i>s</i>	<i>e</i>
<i>D</i>	<i>K</i>	<i>R</i>	<i>N</i>	<i>S</i>	<i>H</i>	<i>G</i>	<i>N</i>	<i>F</i>	<i>I</i>		<i>n</i>	<i>c</i>	<i>e</i>	<i>o</i>	<i>f</i>	<i>l</i>	<i>i</i>	<i>g</i>	<i>h</i>	<i>t</i>
<i>V</i>	<i>J</i>	<i>Y</i>	<i>Q</i>	<i>T</i>	<i>Q</i>	<i>U</i>	<i>X</i>	<i>Q</i>	<i>B</i>	⇒	<i>l</i>	<i>i</i>	<i>e</i>	<i>s</i>	<i>t</i>	<i>h</i>	<i>e</i>	<i>n</i>	<i>u</i>	<i>a</i>
<i>Q</i>	<i>V</i>	<i>Y</i>	<i>U</i>	<i>V</i>	<i>L</i>	<i>L</i>	<i>T</i>	<i>R</i>	<i>E</i>		<i>n</i>	<i>c</i>	<i>e</i>	<i>o</i>	<i>f</i>	<i>i</i>	<i>q</i>	<i>l</i>	<i>u</i>	<i>s</i>
<i>V</i>	<i>J</i>	<i>Y</i>	<i>Q</i>	<i>T</i>	<i>M</i>	<i>K</i>	<i>Y</i>	<i>R</i>	<i>D</i>		<i>i</i>	<i>o</i>	<i>n</i>							
<i>M</i>	<i>F</i>	<i>D</i>																		

So, if we decrypt K1 with the keyword *PALIMPSEST* we find that the first part reads: *between subtle shading and the absence of light lies the nuance of illusion.*

The misspelling of illusion was intentional, according to Sanborn. *When I asked about the misspellings and asked if they were accidental or deliberate, Sanborn said that they were deliberate, but it was less important \*what\* they were. He said, and I quote: “it’s more the orientation of those letters that’s useful there.” Later on in the evening he repeated that point, saying it was the “positioning” that was important.* -Post by Elonka Dunin to the Kryptos group [4].

### 3.4 The Second Section of Kryptos

Sanborn used a Vigenère cipher with ***ABSCISSA*** as the keyword for the second part of Kryptos (abbreviated K2). Abscissa by definition is the horizontal coordinate of a point in a plane Cartesian coordinate system obtained by measuring parallel to the x-axis [9, p.4].

Notice below how the sculpture's ciphertext becomes plaintext by applying the key ***ABSCISSA*** using the Kryptos sculpture's Vigenère tableau in Table 3.0.2.

$\Rightarrow$

The plaintext to the second section on the sculpture states:

*it was totally invisible hows that possible ? they used the earths magnetic field x the information was gathered and transmitted undergruund to an unknown location x does langley know about this ? they should its buried out there somewhere x who knows the exact location ? only ww this was his last message x thirty eight degrees fifty seven minutes six point five seconds north seventy seven degrees eight minutes forty four seconds west id by rows*

Notice how the plaintext ends with *west id by rows*. In 2006 Jim Sanborn announced that an error existed and that a letter had been left out of the ciphertext shown on the sculpture [23]. The Kryptos group (an online chat group found at <http://tech.groups.yahoo.com/group/kryptos/>) noticed that by adding an *x* in the plaintext between the **W** and the **E** in the ciphertext they could have the message *westxlayertwo*, which Sanborn approved as the correct plaintext message. The *x* in the plaintext resembles a period or a break in the text and when Sanborn removed it (for artistic reasons [20]), it altered the decryption.

The second section references to the former CIA director, William Webster (*only ww*); Sanborn has admitted to giving William Webster the key to deciphering the sculpture's message [16]. The second message also includes latitude and longitude directions that point to a place within the agency, *who knows the exact location?*. The exact location varies upon which mapping system is used, though it is somewhere near the sculpture itself: *thirty eight degrees fifty seven minutes six point five seconds north seventy seven degrees eight minutes forty four seconds west*.

In 2005, Sanborn told *Wired News*, "In part of the code that's been deciphered, I refer to an act that took place when I was at the agency and a location that's on the ground of the agency...So in order to find that place, you have to decipher the piece and then go to the agency and find that place [22]." This quote may help us later on in the deciphering process.



## CHAPTER 4

### TRANSPOSITIONS

A transposition cipher is one in which the plaintext letters themselves are not changed but the order in which they appear is altered. You may see in the newspaper, unscramble the letters to reveal the hidden word. This is a simple form of a transposition.

**Example 4.1.** *Unscramble the following*

***YCGROYPOTL***

It would not take long to figure out that the hidden message shown above was the word *cryptology*.

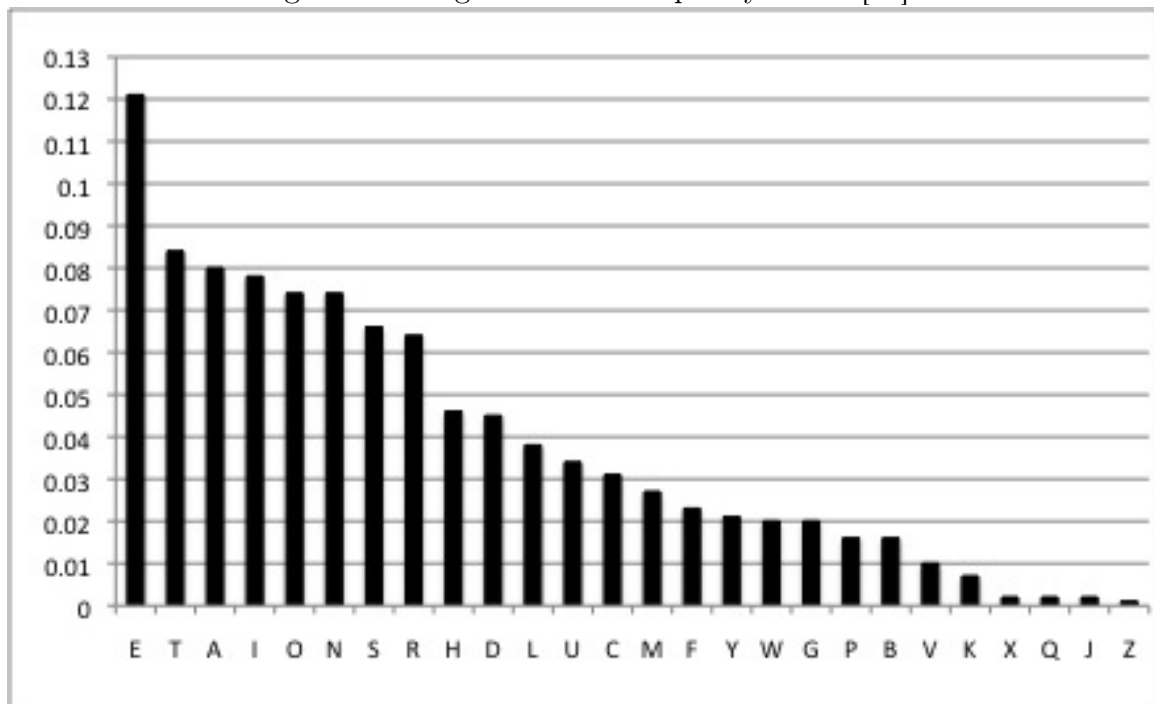
If the ciphertext exhibits a frequency distribution very similar to the English language, it is most likely a transposition. This can then often be attacked by anagramming, moving pieces of ciphertext around and then looking for sections that look like anagrams of English words. Once such anagrams have been found, they reveal information about the transposition pattern and can as a result be extended to form words and then sentences. Since transposition does not affect the frequency of individual letters, transpositions can easily be detected by the cryptanalyst by doing a frequency count.

The sample size of text greatly affects the outcome of the statistical results. If a smaller sample size is chosen, the frequency count will be less-accurate than a frequency count taken from a larger sample of English text. Therefore, if the frequencies of the letters in a text are different than those that appear in Table 4.0.1 below, it does not mean that the text is not English. However, the variance of the percentage of vowels (a,e,i,o,u) appearing in English text is quite small, so they can be used as a reliable test even with small sample sizes. Table 4.0.1 is read left to right, starting with the letter with the highest frequency and descends in order of letter frequencies.

Table 4.0.1. English Letter Frequency [13]

<i>E</i> .121	<i>T</i> .084	<i>A</i> .080	<i>I</i> .078	<i>O</i> .074	<i>N</i> .074	<i>S</i> .066	<i>R</i> .064
<i>H</i> .046	<i>D</i> .045	<i>L</i> .038	<i>U</i> .034	<i>C</i> .031	<i>M</i> .027	<i>F</i> .023	<i>Y</i> .021
<i>W</i> .020	<i>G</i> .020	<i>P</i> .016	<i>B</i> .016	<i>V</i> .010	<i>K</i> .007	<i>X</i> .002	<i>Q</i> .002
<i>J</i> .002	<i>Z</i> .001						

Figure 4.1. English Letter Frequency Chart [13]



The data above were generated from the following texts (most of these literary works were obtained from Project Gutenberg): “The Adventures of Sherlock Holmes,” Arthur Conan Doyle; “Alice’s Adventures in Wonderland,” Lewis Carroll; William Jefferson Clinton’s Inaugural Presidential Address, given from noon to 12:15 P.M., January 20, 1993; “Crime And Punishment,” Fyodor Dostoevsky; “Around the World in 80 Days,” Jules Verne; “Flat-land,” Edwin A. Abbott; “Frankenstein,” Mary W. Shelley; “Moby Dick or the Whale,” Herman Melville; “A Tale of Two Cities,” Charles Dickens; “Up From Slavery,” Booker Taliaferro Washington; “The Voyage Out,” Virginia Woolf; “Pride and Prejudice,” Jane

Austen; “Ulysses,” James Joyce; “A Book of Natural History,” by Various; “The War of the Worlds,” H. G. Wells; “New York Times Current History: The European War, Vol. 2, No. 1,” April, 1915, by Various; “Little Journeys to the Homes of the Great,” Vol. 1 of 14, Elbert Hubbard; “History of the United States,” Charles A. Beard and Mary R. Beard; “Custom and Myth,” Andrew Lang; “English Literature,” William J. Long; “The Symbolism of Freemasonry,” Albert G. Mackey; “Roget’s Thesaurus of English Words and Phrases,” Roget; Full text of the Kyoto protocol on climate change; and Wikipedia articles on American Football, History of American Football, Space Shuttle, Tropical Cyclone, Styracosaurus, Sunspots, Photography, Bill James, Dolomite, Hollywood Pop Academy, and Sabermetrics. In total, these texts constitute approximately 3.3 million words of English text.

*Digrams* are pairs of characters. There are  $26 \times 26 = 676$  possible digrams, but not all of them occur in the English language. Some commonly used digrams are ‘th’, ‘in’, and ‘he.’ *Trigrams* are triples of characters. Few trigrams occur often, but there are  $26 \times 26 \times 26 = 17,576$  possibilities. Trigrams that most often appear in texts include, ‘the’, ‘are’, ‘and’, ‘for’, et cetera. Paul Garrett researched frequencies of digrams and trigrams in his book *Making Breaking Codes: An Introduction to Cryptology*. He looked at approximately one megabyte of old email after removing the headers and found the approximate frequencies shown in Tables 4.0.2, 4.0.3, 4.0.4, and 4.0.5 [8, p. 35,36]. In English, a frequency analysis counting the letters of the alphabet along with blanks reveals that blanks are more common, representing 17%-18% of the messages [8, p. 32]. Therefore if blanks were encrypted as a letter, their high frequency could give a hint on the decryption. For this reason blanks are normally removed from messages being encrypted.

Table 4.0.2. The top 77 digrams occurring within words in the sample, with percentages (including blanks)

<i>th</i> 3.81	<i>in</i> 2.59	<i>he</i> 2.17	<i>er</i> 1.95	<i>re</i> 1.85	<i>on</i> 1.59	<i>an</i> 1.59
<i>at</i> 1.54	<i>on</i> 1.43	<i>or</i> 1.26	<i>es</i> 1.26	<i>ha</i> 1.24	<i>to</i> 1.22	<i>te</i> 1.21
<i>is</i> 1.18	<i>ti</i> 1.17	<i>it</i> 1.16	<i>en</i> 1.13	<i>nt</i> 1.09	<i>ng</i> 1.08	<i>al</i> 1.07
<i>se</i> 1.05	<i>st</i> 1.01	<i>nd</i> 0.98	<i>le</i> 0.91	<i>ar</i> 0.90	<i>me</i> 0.90	<i>hi</i> 0.86
<i>ve</i> 0.85	<i>of</i> 0.84	<i>ed</i> 0.78	<i>co</i> 0.74	<i>as</i> 0.73	<i>ll</i> 0.72	<i>ne</i> 0.70
<i>om</i> 0.70	<i>ri</i> 0.68	<i>ic</i> 0.67	<i>ro</i> 0.67	<i>ea</i> 0.66	<i>et</i> 0.64	<i>ur</i> 0.64
<i>io</i> 0.64	<i>ra</i> 0.62	<i>li</i> 0.62	<i>no</i> 0.62	<i>so</i> 0.62	<i>be</i> 0.61	<i>de</i> 0.59
<i>ma</i> 0.59	<i>si</i> 0.58	<i>ly</i> 0.54	<i>ut</i> 0.53	<i>ot</i> 0.53	<i>pr</i> 0.53	<i>fo</i> 0.53
<i>yo</i> 0.52	<i>il</i> 0.50	<i>ca</i> 0.50	<i>pe</i> 0.50	<i>ch</i> 0.49	<i>ho</i> 0.49	<i>ul</i> 0.47
<i>ce</i> 0.47	<i>ta</i> 0.45	<i>di</i> 0.45	<i>rs</i> 0.45	<i>el</i> 0.44	<i>ge</i> 0.44	<i>us</i> 0.44
<i>ec</i> 0.42	<i>ss</i> 0.42	<i>ac</i> 0.41	<i>ct</i> 0.41	<i>em</i> 0.41	<i>wh</i> 0.40	<i>oo</i> 0.40

Table 4.0.3. The top 77 digrams occurring within words in the sample, with percentages (without blanks)

<i>th</i> 2.63	<i>in</i> 2.08	<i>he</i> 1.75	<i>er</i> 1.67	<i>re</i> 1.52	<i>on</i> 1.33	<i>es</i> 1.32
<i>an</i> 1.29	<i>at</i> 1.28	<i>ti</i> 1.26	<i>nt</i> 1.16	<i>ou</i> 1.16	<i>to</i> 1.13	<i>st</i> 1.12
<i>ha</i> 1.05	<i>or</i> 1.05	<i>et</i> 1.03	<i>en</i> 1.01	<i>te</i> 1.01	<i>is</i> 0.98	<i>it</i> 0.97
<i>ea</i> 0.93	<i>se</i> 0.90	<i>al</i> 0.89	<i>ng</i> 0.89	<i>nd</i> 0.81	<i>ed</i> 0.76	<i>hi</i> 0.75
<i>le</i> 0.75	<i>ar</i> 0.74	<i>si</i> 0.73	<i>me</i> 0.73	<i>so</i> 0.71	<i>of</i> 0.70	<i>ve</i> 0.68
<i>ri</i> 0.64	<i>as</i> 0.64	<i>om</i> 0.64	<i>ra</i> 0.61	<i>no</i> 0.61	<i>ne</i> 0.60	<i>co</i> 0.60
<i>ro</i> 0.59	<i>ll</i> 0.59	<i>ta</i> 0.58	<i>ic</i> 0.57	<i>ot</i> 0.57	<i>tt</i> 0.57	<i>li</i> 0.57
<i>yo</i> 0.52	<i>ur</i> 0.51	<i>ec</i> 0.51	<i>io</i> 0.51	<i>de</i> 0.51	<i>di</i> 0.51	<i>ma</i> 0.51
<i>ei</i> 0.49	<i>be</i> 0.49	<i>sa</i> 0.47	<i>ss</i> 0.47	<i>el</i> 0.46	<i>em</i> 0.46	<i>rs</i> 0.45
<i>fo</i> 0.44	<i>ut</i> 0.44	<i>ly</i> 0.44	<i>rt</i> 0.43	<i>ca</i> 0.42	<i>pr</i> 0.42	<i>na</i> 0.42
<i>ts</i> 0.41	<i>ho</i> 0.41	<i>il</i> 0.41	<i>pe</i> 0.40	<i>ch</i> 0.40	<i>ul</i> 0.38	<i>ee</i> 0.38

Table 4.0.4. The top 77 most common trigrams within English words, with percentages (including blanks)

<i>the</i> 2.44	<i>ing</i> 1.26	<i>and</i> 0.82	<i>hat</i> 0.78	<i>tha</i> 0.77	<i>ion</i> 0.75	<i>you</i> 0.67
<i>ent</i> 0.66	<i>for</i> 0.63	<i>tio</i> 0.63	<i>thi</i> 0.60	<i>her</i> 0.51	<i>ati</i> 0.47	<i>our</i> 0.47
<i>ere</i> 0.45	<i>all</i> 0.43	<i>ter</i> 0.43	<i>ver</i> 0.40	<i>not</i> 0.40	<i>hin</i> 0.40	<i>ome</i> 0.36
<i>oul</i> 0.36	<i>uld</i> 0.36	<i>int</i> 0.34	<i>rea</i> 0.34	<i>pro</i> 0.34	<i>res</i> 0.33	<i>ate</i> 0.33
<i>hav</i> 0.30	<i>ave</i> 0.30	<i>ill</i> 0.30	<i>his</i> 0.30	<i>com</i> 0.30	<i>ons</i> 0.30	<i>are</i> 0.28
<i>ple</i> 0.28	<i>ers</i> 0.28	<i>con</i> 0.27	<i>ess</i> 0.27	<i>out</i> 0.27	<i>one</i> 0.26	<i>ith</i> 0.25
<i>som</i> 0.25	<i>ive</i> 0.25	<i>tin</i> 0.25	<i>nce</i> 0.24	<i>ble</i> 0.24	<i>ted</i> 0.24	<i>han</i> 0.23
<i>ine</i> 0.23	<i>per</i> 0.23	<i>ect</i> 0.23	<i>nte</i> 0.23	<i>wit</i> 0.22	<i>men</i> 0.22	<i>but</i> 0.22
<i>wou</i> 0.21	<i>ica</i> 0.21	<i>eve</i> 0.21	<i>cal</i> 0.21	<i>pre</i> 0.21	<i>cou</i> 0.21	<i>lin</i> 0.21
<i>est</i> 0.20	<i>eri</i> 0.21	<i>mor</i> 0.20	<i>ser</i> 0.20	<i>ore</i> 0.19	<i>any</i> 0.19	<i>abi</i> 0.19
<i>tic</i> 0.19	<i>urs</i> 0.19	<i>ant</i> 0.19	<i>sti</i> 0.18	<i>ear</i> 0.18	<i>hou</i> 0.18	<i>ies</i> 0.18

Table 4.0.5. The top 77 most common trigrams within English words, with percentages (without blanks)

<i>the</i> 1.49	<i>ing</i> 0.77	<i>tha</i> 0.52	<i>and</i> 0.50	<i>hat</i> 0.47	<i>ion</i> 0.45	<i>ent</i> 0.43
<i>you</i> 0.41	<i>thi</i> 0.38	<i>for</i> 0.38	<i>ati</i> 0.38	<i>tio</i> 0.38	<i>her</i> 0.35	<i>ere</i> 0.35
<i>eth</i> 0.34	<i>int</i> 0.32	<i>our</i> 0.28	<i>tth</i> 0.27	<i>all</i> 0.27	<i>rea</i> 0.26	<i>ter</i> 0.26
<i>nth</i> 0.26	<i>ome</i> 0.25	<i>hin</i> 0.25	<i>ver</i> 0.25	<i>not</i> 0.24	<i>res</i> 0.23	<i>est</i> 0.22
<i>oul</i> 0.22	<i>ont</i> 0.22	<i>ate</i> 0.21	<i>uld</i> 0.21	<i>ers</i> 0.21	<i>tin</i> 0.21	<i>oth</i> 0.20
<i>pro</i> 0.20	<i>sth</i> 0.20	<i>ons</i> 0.20	<i>his</i> 0.19	<i>ith</i> 0.19	<i>ave</i> 0.19	<i>eri</i> 0.19
<i>sin</i> 0.19	<i>ess</i> 0.18	<i>are</i> 0.18	<i>hav</i> 0.18	<i>ist</i> 0.18	<i>ill</i> 0.18	<i>out</i> 0.18
<i>com</i> 0.18	<i>rth</i> 0.18	<i>ese</i> 0.17	<i>ore</i> 0.17	<i>ple</i> 0.17	<i>con</i> 0.17	<i>one</i> 0.16
<i>att</i> 0.16	<i>iti</i> 0.16	<i>ert</i> 0.16	<i>ica</i> 0.16	<i>ein</i> 0.16	<i>eto</i> 0.16	<i>som</i> 0.16
<i>han</i> 0.15	<i>oft</i> 0.15	<i>nte</i> 0.15	<i>ine</i> 0.15	<i>sto</i> 0.15	<i>ted</i> 0.15	<i>ive</i> 0.15
<i>ear</i> 0.15	<i>fth</i> 0.15	<i>nce</i> 0.15	<i>ret</i> 0.14	<i>ngt</i> 0.14	<i>ble</i> 0.14	<i>lin</i> 0.14

#### 4.1 Columnar Transpositions

There are many different types of transpositions. For example, there are *block transpositions* where text is moved around within blocks, or *columnar transpositions* where the text is rearranged by columns.

**Example 4.2.** Use a columnar transposition to encipher the plaintext: *we hold these truths to be self evident that all men are created equal.*

First we pick a keyword, say **ALGEBRA**. Second we number each letter (column) in ascending order of the alphabet ( $A=1, B=2, \dots, Z=26$ ). Third, we write our plaintext out in rows with length of the keyword.

<i>A</i>	<i>L</i>	<i>G</i>	<i>E</i>	<i>B</i>	<i>R</i>	<i>A</i>
<i>1</i>	<i>6</i>	<i>5</i>	<i>4</i>	<i>3</i>	<i>7</i>	<i>2</i>
<i>W</i>	<i>E</i>	<i>H</i>	<i>O</i>	<i>L</i>	<i>D</i>	<i>T</i>
<i>H</i>	<i>E</i>	<i>S</i>	<i>E</i>	<i>T</i>	<i>R</i>	<i>U</i>
<i>T</i>	<i>H</i>	<i>S</i>	<i>T</i>	<i>O</i>	<i>B</i>	<i>E</i>
<i>S</i>	<i>E</i>	<i>L</i>	<i>F</i>	<i>E</i>	<i>V</i>	<i>I</i>
<i>D</i>	<i>E</i>	<i>N</i>	<i>T</i>	<i>T</i>	<i>H</i>	<i>A</i>
<i>T</i>	<i>A</i>	<i>L</i>	<i>L</i>	<i>M</i>	<i>E</i>	<i>N</i>
<i>A</i>	<i>R</i>	<i>E</i>	<i>C</i>	<i>R</i>	<i>E</i>	<i>A</i>
<i>T</i>	<i>E</i>	<i>D</i>	<i>E</i>	<i>Q</i>	<i>U</i>	<i>A</i>
<i>L</i>						

Now, (our fourth step) transposing our plaintext in ascending order of the columns we have the ciphertext:

*Column 1*
*Column 2*
*Column 3*
*Column 4*
*Column 5*
*Column 6*
*Column 7*

*WHTSDTATLTUEIANAALTOETMRQOETFTLCEHSSLNLEDEEHEEAREDRBVHEEU*

Looking for anagrams throughout the plaintext can lead a decipherer in the right direction. That is, guessing probable column lengths and shifting around to find anagrams will lead to revealing the plaintext. This leads us to a stronger transposition, the *double columnar transposition*.

The double columnar transposition is the columnar transposition in the previous example, completed twice. It can be a strong cipher when the key is long and the

columns are short. The transposition can be made stronger the more transpositions we do along with using different keywords for each transposition. This is because it makes looking for correct anagrams in the decryption process more difficult. The double columnar transposition was used by the United States Army [7, p. 54]. It was a quick method to encipher a message and, with knowledge of the key, to decipher. With a high probability that their message would be intercepted their main concern was that it would take the interceptor a great length of time, maybe hours or days, to decrypt the message without knowing the key. Thus by the time the interceptor figured out the plaintext message, the information would no longer be valuable. Therefore, the double columnar transposition proved useful [7, p. 37-55].

**Example 4.3.** *We will perform a second columnar transposition using Example 4.2's columnar transposition. The keyword will now be **ABELIAN**.*

<i>A</i>	<i>B</i>	<i>E</i>	<i>L</i>	<i>I</i>	<i>A</i>	<i>N</i>
<i>1</i>	<i>3</i>	<i>4</i>	<i>6</i>	<i>5</i>	<i>2</i>	<i>7</i>
<i>W</i>	<i>H</i>	<i>T</i>	<i>S</i>	<i>D</i>	<i>T</i>	<i>A</i>
<i>T</i>	<i>L</i>	<i>T</i>	<i>U</i>	<i>E</i>	<i>I</i>	<i>A</i>
<i>N</i>	<i>A</i>	<i>A</i>	<i>L</i>	<i>T</i>	<i>O</i>	<i>E</i>
<i>T</i>	<i>M</i>	<i>R</i>	<i>Q</i>	<i>O</i>	<i>E</i>	<i>T</i>
<i>F</i>	<i>T</i>	<i>L</i>	<i>C</i>	<i>E</i>	<i>H</i>	<i>S</i>
<i>S</i>	<i>L</i>	<i>N</i>	<i>L</i>	<i>E</i>	<i>D</i>	<i>E</i>
<i>E</i>	<i>H</i>	<i>E</i>	<i>E</i>	<i>A</i>	<i>R</i>	<i>E</i>
<i>D</i>	<i>R</i>	<i>B</i>	<i>V</i>	<i>H</i>	<i>E</i>	<i>E</i>
<i>U</i>						

Therefore our ciphertext will now read:

Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7
<i>W</i>	<i>T</i>	<i>N</i>	<i>T</i>	<i>F</i>	<i>S</i>	<i>E</i>
<i>D</i>	<i>U</i>	<i>T</i>	<i>I</i>	<i>O</i>	<i>E</i>	<i>H</i>
<i>R</i>	<i>E</i>	<i>H</i>	<i>D</i>	<i>R</i>	<i>E</i>	<i>H</i>
<i>L</i>	<i>A</i>	<i>M</i>	<i>T</i>	<i>L</i>	<i>H</i>	<i>R</i>
<i>T</i>	<i>T</i>	<i>A</i>	<i>R</i>	<i>L</i>	<i>N</i>	<i>E</i>
<i>B</i>	<i>D</i>	<i>E</i>	<i>T</i>	<i>O</i>	<i>E</i>	<i>E</i>
<i>A</i>	<i>H</i>	<i>S</i>	<i>U</i>	<i>L</i>	<i>Q</i>	<i>C</i>
<i>L</i>	<i>E</i>	<i>V</i>	<i>A</i>	<i>A</i>	<i>E</i>	<i>T</i>
<i>S</i>	<i>E</i>	<i>E</i>	<i>E</i>	<i>E</i>	<i>E</i>	<i>E</i>

## 4.2 The Third Section of Kryptos

The third part of Kryptos (K3) was coded with a transposition cipher. It reads:  
*slowly desperatly slowly the remains of passage debris*  
*that encumbered the lower part of the doorway was removed*

*with trembling hands i made a tiny breach in the upper left hand corner and then widening the hole a little i inserted the candle and peered in the hot air escaping from the chamber caused the flame to flicker but presently details of the room within emerged from the mist x can you see anything q?*

This is a paraphrase of the archaeologist Howard Carter's description of his opening of King Tut's tomb in his 1923 book, *The Tomb of Tutankhamen* [1]. Sanborn included this message because he said it has inspired him since he was a child [15]. *"It was probably the most exciting thing I ever read as a child... I started out life as an archeologist, thought that that was what I wanted to be. (It) sounded very exciting to discover things, whether it was an arrowhead, or a shark's tooth on a beach, or a dinosaur bone, or an artifact... and that moment of discovery was very exciting... and I think that that same moment of discovery is what a cryptographer feels when they decode a message that's been hidden for years... and so I chose that statement (Howard Carter's account), to demonstrate, or really give the essence, or the feeling of discovery [19]."*

One method possibly used to encrypt the text was a double columnar transposition as follows: write the plaintext into a 28 columns.

Figure 4.2. Section 3 Plaintext Written into 28 Columns

S	L	O	W	L	Y	D	E	S	P	E	R	A	T	L	Y	S	L	O	W	L	Y	T	H	E	R	E	M
A	I	N	S	O	F	P	A	S	S	A	G	E	D	E	B	R	I	S	T	H	A	T	E	N	C	U	M
B	E	R	E	D	T	H	E	L	O	W	E	R	P	A	R	T	O	F	T	H	E	D	O	O	R	W	A
Y	W	A	S	R	E	M	O	V	E	D	W	I	T	H	T	R	E	M	B	L	I	N	G	H	A	N	D
S	I	M	A	S	E	A	T	I	N	Y	B	R	E	A	C	H	I	N	T	H	E	U	P	P	E	R	L
E	F	T	H	A	N	D	C	O	R	N	E	R	A	N	D	T	H	E	N	W	I	D	E	N	I	N	G
T	H	E	H	O	L	E	A	L	I	T	T	L	E	I	I	N	S	E	R	T	E	D	T	H	E	C	A
N	D	L	E	A	N	D	P	E	E	R	E	D	I	N	T	H	E	H	O	T	A	I	R	E	S	C	A
P	I	N	G	F	R	O	M	T	H	E	C	H	A	M	B	E	R	C	A	U	S	E	D	T	H	E	F
L	A	M	E	T	O	F	L	I	V	K	E	R	B	U	T	P	R	E	S	E	N	T	L	Y	D	E	T
A	I	L	S	O	F	T	H	E	R	O	O	M	W	I	T	H	I	N	E	M	E	R	G	E	D	F	R
O	M	T	H	E	M	I	S	T	X	C	A	N	Y	O	U	S	E	E	A	N	Y	T	H	I	N	G	Q

Read off by columns, from last to first, and write a new grid with 21 columns.



Figure 4.3. Section 3 Plaintext New Grid with 21 Columns

M	M	A	D	L	G	A	A	F	T	R	Q	E	U	W	N	R	N	C	C	E
E	F	G	R	C	R	A	E	I	E	S	H	D	D	N	E	N	O	H	P	N
H	E	T	Y	E	I	H	E	O	G	P	E	T	R	D	L	G	H	T	T	D
N	U	D	D	I	E	T	R	T	Y	A	E	I	E	I	E	A	S	N	E	Y
L	H	H	L	H	W	T	T	U	E	M	N	W	T	T	B	T	N	R	O	A
S	E	A	O	S	F	M	N	E	E	H	C	E	N	E	L	I	O	E	I	H
S	E	R	R	I	E	S	R	T	R	H	T	N	H	E	P	H	S	Y	B	R
T	C	D	I	T	B	T	T	U	L	E	A	H	A	N	I	N	M	U	I	O
T	D	P	T	E	A	E	I	A	B	W	Y	A	E	R	I	R	R	L	D	H
R	M	N	R	G	E	W	B	E	T	E	C	E	O	A	A	A	W	D	Y	N
T	R	E	K	O	C	P	S	O	E	N	R	I	E	H	C	R	X	S	S	L
V	I	O	L	E	T	I	E	T	E	A	E	O	T	C	A	P	M	L	H	S
D	P	H	M	A	D	E	D	O	F	T	I	Y	F	T	E	E	N	L	N	R
O	F	M	L	O	D	R	D	A	O	A	F	T	O	E	W	S	E	S	A	H
H	E	G	E	S	H	O	N	R	A	M	T	E	L	N	M	L	T	L	I	E
W	I	F	H	D	I	A	I	M	S	A	B	Y	S	E	T	N	P	L	A	O

Repeat this process, writing the last column into the first row making the ciphertext we see on the sculpture.

Figure 4.4. Section 3 Plaintext Transposed to Ciphertext

E	N	D	Y	A	H	R	O	H	N	L	S	R	H	E	O	C	P	T	E	O	I	B	I	D	Y	S	H	N	A	I	A	
C	H	T	N	R	E	Y	U	L	D	S	L	L	S	L	L	N	O	H	S	N	O	S	M	R	W	X	M	N	E			
T	P	R	N	G	A	T	I	H	N	R	A	R	P	E	S	L	N	N	E	L	E	B	L	P	I	I	A	C	A	E		
W	M	T	W	N	D	I	T	E	E	N	R	A	H	C	T	E	N	E	U	D	R	E	T	N	H	A	E	O	E			
T	F	O	L	S	E	D	T	I	W	E	N	H	A	E	I	O	Y	T	E	Y	Q	H	U	E	E	N	C	T	A	Y	C	R
E	I	F	T	B	R	S	P	A	M	H	H	E	W	E	N	A	T	A	M	A	T	E	G	Y	E	E	R	L	B			
T	E	E	F	O	A	S	F	I	O	T	U	E	T	U	A	E	O	T	O	A	R	M	A	E	E	R	T	N	R	T	I	
B	S	E	D	D	N	I	A	A	H	T	T	M	S	T	E	W	P	I	E	R	O	A	G	R	I	E	W	F	E	B		
A	E	C	T	D	D	H	I	L	C	E	I	H	S	I	T	E	G	O	E	A	O	S	D	D	R	Y	D	L	O	R	I	T
R	K	L	M	L	E	H	A	G	T	D	H	A	R	D	P	N	E	O	H	M	G	F	M	F	E	U	H	E				
E	C	D	M	R	I	P	F	E	I	M	E	H	N	L	S	S	T	T	R	T	V	D	O	H	W							

The decryption process proceeds the same way, in reverse.

## CHAPTER 5

### THE FOURTH SECTION OF KRYPTOS

The fourth encrypted text (abbreviated as K4) on the statue reads:

OBKR  
UOXOGHULBSOLIFBBWFLRVQQPRNGKSSO  
TWTQSJQSSEKZZWATJKLUDIAWINFBNYP  
VTTMZFPKWGDKZXTJCDIGKUHUAUEKCAR

To date, the fourth section has still not been decrypted. It has been a mystery for over 20 years. This seems quite strange that an artist, not even a cryptologist, can create a code that is difficult for some of the worlds greatest cryptologists. Why haven't cryptologists been able to solve this mystery?

Some possible reasons are:

- It is shorter than any but the first. Longer strings of ciphertext are easier to decode because we can look for patterns.
- We don't know what cipher he used and the ciphertext displays no periodicity. There are many ways to encipher a message, without knowing which way Sanborn chose, testing every possibility is tedious.
- Sanborn may have made mistakes in enciphering the message. A cryptographer somewhere may have already tried the right key but the translation went unnoticed because of propagated errors in the ciphertext.
- Ways he could have also concealed the plaintext before encrypting include:
  - Spelled words incorrectly,
  - Padded with nulls

#### 5.1 Clues

In interviews with Sanborn the only clue he verbally gives to decode the fourth part is [while pointing at the statue] stating all you need is "in plain sight" [12]. Some cryptologists believe that there are many hints embedded within the

sculpture. For example, some of the letters cut into the copper are slightly higher than others in the same row. Why is this? We also note that all of the letters are cut all the way through the copper, so that sunlight shines through the sculpture to create interesting patterns of light and shadow on the ground. Do these patterns provide a clue to cracking the code? The directions given in K2 could be guiding us somewhere, or why is there an extra L in one of the rows of lettering on the statue? The previous misspellings *iglusion* and *undergruund* were supposedly intentional. Further, the copper scroll is only the main part of Sanborn's work. There are several other accompanying objects scattered around the CIA grounds, including stone/copper slabs with messages like SOS, LUCID MEMORY, T IS YOUR POSITION, SHADOW FORCES, VIRTUALLY INVISIBLE, DIGETAL INTERPRETATU, and RQ (or YR if you look from the other direction) written in Morse code. There is also a stone with a compass pointing north.

Sanborn provided *The New York Times* in 2010 with information about the fourth part [15]. He announced that the 64th through 69th characters of the ciphertext appearing as **NYPVTT** are decrypted as *berlin*. The hint *berlin* could be just to show that there is a message; that it really is capable of being decoded. Sanborn has been harassed for the last 20 years by amateurs asking him what is the key and blaming him that there must not be one, there is no message. Another thought is that decoding the text has to do with Enigma. Or maybe it can be decoded with a German method or it is a quote from something or someone from Germany, since Berlin is in the plaintext. Whatever Sanborn's reasonings were to announce the hint, we will not know until the fourth part has been completely solved.

## 5.2 ABCD

The Vigenère tableau appearing on one side of the sculpture, (see Table 1.0.1), has four extra columns starting with *ABCD*. Maybe these are not extra at all. Maybe Sanborn placed it there because we were indeed supposed to use these columns.

If in fact the encryption method was Vigenère we would be able to tell by the index of coincidence for possible key lengths. If the index of coincidence is close to .0667 it could possibly resemble English text with a Vigenère cipher. However, looking at Table 5.2.1 below, we can see that all plausible key lengths seem to be

unlikely (close to a .0385 index of coincidence).

Table 5.2.1. Index of Coincidence for Possible Key Lengths

key length	1	2	3	4	5	6
IC	0.0361	0.0352	0.0297	0.0328	0.0392	0.0301
key length	7	8	9	10	11	12
IC	0.0419	0.0262	0.0256	0.0433	0.0440	0.023

Therefore we can conclude that it is highly unlikely that the fourth section of Kryptos was encrypted with a Vigenère cipher.

Another similar possibility is that the method of encryption was Vigenère with key **ABCD** followed by a transposition. This would leave little or no trace of periodicity in the ciphertext, but we can still test for the hypothesis as follows.

If a plaintext is encrypted with **ABCD** we can find the frequency with which each letter is expected to appear in the ciphertext. Then we can compare the expected frequencies with the actual frequencies to see whether this hypothesis is correct.

Let  $\phi(\mathbf{p}, \mathbf{K})$  denote the ciphertext character obtained from the Kryptos tableau (see Table 3.0.2) with plaintext character  $\mathbf{p}$  and key character  $\mathbf{K}$ . Key letters of **A**, **B**, **C**, and **D** all have an equal probability of .25 since this is the repeated key being considered. Thus, we have the equation below to find the probability of some ciphertext letter  $\alpha$ .

$$\begin{aligned}
 \text{Probability}(\text{ciphertext } \alpha) &= \sum_{\substack{\mathbf{p}, \mathbf{K} \\ \phi(\mathbf{p}, \mathbf{K}) = \alpha}} \text{Probability}(\mathbf{p}) \times \text{Probability}(\mathbf{K}) \\
 &= \sum_{\substack{\mathbf{p}, \mathbf{K} \\ \phi(\mathbf{p}, \mathbf{K}) = \alpha}} \text{Probability}(\text{English } \mathbf{p}) \times .25
 \end{aligned}$$

Thus if a plaintext is encrypted with **ABCD**, how can a ciphertext **A** appear using the Kryptos tableau? Taking  $\phi(\mathbf{p}, \mathbf{K}) = \mathbf{A}$  we find that:

$$\phi(\mathbf{k}, \mathbf{A}) = \mathbf{A}; \phi(\mathbf{z}, \mathbf{B}) = \mathbf{A}; \phi(\mathbf{x}, \mathbf{C}) = \mathbf{A}; \phi(\mathbf{w}, \mathbf{D}) = \mathbf{A}.$$

Now looking at the frequency of each letter we can find the expected frequency of **A** in the ciphertext by the equation above. That is, in the plaintext we would expect

that **k** has a frequency of .007, **z** has a frequency of .001, **x** has a frequency of .002, and **w** has a frequency of .02. **A**, **B**, **C**, and **D** all have an equal probability of .25 since this is the repeated key. Therefore we have the expected frequency of **A** to be:

$$(.007)(.25) + (.001)(.25) + (.002)(.25) + (.02)(.25) = .0075.$$

Since there are 97 characters in K4, we would expect about  $(97)(.0075) = .72750$  **A**'s to appear in the ciphertext.

If we have  $\phi(\mathbf{p}, \mathbf{K})=\mathbf{B}$  then we find that:

$$\phi(\mathbf{r}, \mathbf{A})=\mathbf{B}; \phi(\mathbf{k}, \mathbf{B})=\mathbf{B}; \phi(\mathbf{z}, \mathbf{C})=\mathbf{B}; \phi(\mathbf{x}, \mathbf{D})=\mathbf{B}.$$

So, looking at the frequency of each letter we can find the expected frequency of **B** in the ciphertext. That is, in the plaintext we would expect that **r** has a frequency of .064, **k** has a frequency of .007, **z** has a frequency of .001, and **x** has a frequency of .002. Therefore, the expected frequency of **B** is:

$$(.064)(.25) + (.007)(.25) + (.001)(.25) + (.002)(.25) = .01850.$$

Since there are 97 characters in K4, we would expect about  $(97)(.01850) = 1.79450$  **B**'s to appear in the ciphertext.

If we have  $\phi(\mathbf{p}, \mathbf{K})=\mathbf{C}$  then we find that:

$$\phi(\mathbf{y}, \mathbf{A})=\mathbf{C}; \phi(\mathbf{r}, \mathbf{B})=\mathbf{C}; \phi(\mathbf{k}, \mathbf{C})=\mathbf{C}; \phi(\mathbf{z}, \mathbf{D})=\mathbf{C}.$$

So, looking at the frequency of each letter we can find the expected frequency of **C** in the ciphertext. That is, in the plaintext we would expect that **y** has a frequency of .021, **r** has a frequency of .064, **k** has a frequency of .007, and **z** has a frequency of .001. Therefore, the expected frequency of **C** is:

$$(.021)(.25) + (.064)(.25) + (.007)(.25) + (.001)(.25) = .02325.$$

Since there are 97 characters in K4, we would expect about  $(97)(.02325) = 2.5525$  **C**'s to appear in the ciphertext.

Continuing in this manner we are able to calculate each letter's expected frequency. The table below compares the actual frequency of each letter in the ciphertext with its expected frequency we calculated.

Table 5.2.2. Comparison of Letter Frequencies on the Fourth Section of the Sculpture

Letter	Expected Number of Occurrences	Actual Number of Occurrences
A	.72750	4
B	1.7945	5
C	2.5525	2
D	2.619	3
E	4.48625	2
F	4.72875	4
G	5.82	4
H	7.372	2
I	5.723	4
J	4.68025	3
K	3.41925	8
L	4.171	4
M	5.16525	1
N	5.335	3
O	1.6005	5
P	2.91	3
Q	5.06825	4
R	3.41925	4
S	.80025	6
T	1.6005	6
U	5.0925	6
V	4.04975	2
W	3.5405	5
X	3.977	2
Y	3.32225	1
Z	3.51625	4

Karl Pearson, in 1900, suggested what is known today as the *Chi Squared Test* for testing the likelihood that some given data is a sample from a given distribution [17, p. 715]. The test statistic is the following:

$$\chi^2 = \sum_{all\ cells} \frac{(observed - expected)^2}{expected}.$$

We can use this test to determine the likelihood that our hypothesis is correct. That

is, if a Vigenère cipher with key **ABCD** followed by a transposition was a likely encryption.

Applying the Chi Squared Test to the data previously found above we get that  $\chi^2 \approx 96.142$ .

Now for a few definitions; If we have a test statistic (like the one above), the *significance level* is the smallest level of probability for which the observed data will indicate that the hypothesis should be rejected [17, p. 513]. *Degrees of freedom* are the number of values that are free to vary, they do not depend on other values. We now have our Chi Square statistic ( $\chi^2 \approx 96.142$ ), our predetermined level of significance (0.005), and our degrees of freedom (25). Taking a look at a Chi Square distribution table with 25 degrees of freedom and reading along the row we find our value of 96.142 to be too large for the table (it is not even on the table) [17, p. 850, 851]. This implies that the probability of this shift occurring is smaller than the originally chosen significance level of 0.005 or .5%, so the hypothesis that the 4th section was a Vigenère with Key **ABCD** followed by a transposition is rejected.

### 5.3 Running Key and One Time Pad

A *running key cipher* is similar to a Vigenère cipher; however the key as large as the plaintext is chosen from a text or passage that can be found in a book. Most likely the passage is chosen ahead of time and agreed upon by the sender and receiver. The difference between a regular Vigenère cipher and a running key cipher is that the running key will have a long text as its key. This is not to be confused with the *one time pad* which is a cipher with a uniformly random key which is truly unbreakable if used correctly. In a one time pad the key is the same length as (or longer than) the ciphertext, but chosen uniformly at random. It could be possible that Sanborn used a running key cipher. A quote by Jim Sanborn in an interview with *Wired Magazine* says, “*I will say that I have left instructions in the earlier text that refer to later text* [21].” So it is possible that he may be using earlier text on the sculpture to encrypt the fourth part.

As stated above, Sanborn announced that **NYPVTT**, the 64th through 69th characters in K4 are decrypted as **berlin**. If this is true and the Kryptos tableau was used, then the key for these six letters is **ELYOIE**. **ELY** is an English trigram

that is possibly part of an English word in the plaintext of K4. Sanborn misspelled *desperatly* in the plaintext of K2. If we spell the word correctly and add “desperat” on to the key we would have *DESPERATELYOIE* as the key for the 56th through 69th characters. However this would lead to a plaintext of *ktrstmotberlin* which does not appear to be English. Therefore, we can conclude that “desperat” does not come before the trigram *ELY*. *ELYOIE* has 4 vowels in it which provides the possibility that it is English text. However, transposed English certainly seems like a possibility for a running key here.

We can use the Chi-Squared test to determine the likelihood that our ciphertext is English text encrypted with a permuted English text key; that is, if our ciphertext was enciphered with a running key of transposed English. The probability of a ciphertext letter  $\alpha$  is:

$$\begin{aligned} \text{Probability}(\text{ciphertext } \alpha) &= \sum_{\substack{\mathbf{p}, K \\ \phi(\mathbf{p}, K) = \alpha}} \text{Probability}(\mathbf{p}) \times \text{Probability}(K) \\ &= \sum_{\substack{\mathbf{p}, K \\ \phi(\mathbf{p}, K) = \alpha}} \text{Probability}(\text{English } \mathbf{p}) \times \text{Probability}(\text{English } K). \end{aligned}$$

If a plaintext is encrypted with a standard English alphabet how can a ciphertext **A** appear using the Kryptos tableau (see Table 3.0.2)? So if we have  $\phi(\text{plaintext}, \text{KEY}) = \mathbf{A}$ , we find that:

$$\begin{aligned} \phi(\mathbf{k}, \mathbf{A}) &= \mathbf{A}; \phi(\mathbf{r}, \mathbf{S}) = \mathbf{A}; \phi(\mathbf{y}, \mathbf{O}) = \mathbf{A}; \phi(\mathbf{p}, \mathbf{T}) = \mathbf{A}; \phi(\mathbf{t}, \mathbf{P}) = \mathbf{A}; \phi(\mathbf{o}, \mathbf{Y}) = \mathbf{A}; \\ \phi(\mathbf{s}, \mathbf{R}) &= \mathbf{A}; \phi(\mathbf{a}, \mathbf{K}) = \mathbf{A}; \phi(\mathbf{b}, \mathbf{Z}) = \mathbf{A}; \phi(\mathbf{c}, \mathbf{X}) = \mathbf{A}; \phi(\mathbf{d}, \mathbf{W}) = \mathbf{A}; \phi(\mathbf{e}, \mathbf{V}) = \mathbf{A}; \\ \phi(\mathbf{f}, \mathbf{U}) &= \mathbf{A}; \phi(\mathbf{g}, \mathbf{Q}) = \mathbf{A}; \phi(\mathbf{h}, \mathbf{N}) = \mathbf{A}; \phi(\mathbf{i}, \mathbf{M}) = \mathbf{A}; \phi(\mathbf{j}, \mathbf{L}) = \mathbf{A}; \phi(\mathbf{l}, \mathbf{J}) = \mathbf{A}; \\ \phi(\mathbf{m}, \mathbf{I}) &= \mathbf{A}; \phi(\mathbf{n}, \mathbf{H}) = \mathbf{A}; \phi(\mathbf{q}, \mathbf{G}) = \mathbf{A}; \phi(\mathbf{u}, \mathbf{F}) = \mathbf{A}; \phi(\mathbf{v}, \mathbf{E}) = \mathbf{A}; \phi(\mathbf{w}, \mathbf{D}) = \mathbf{A}; \\ \phi(\mathbf{x}, \mathbf{C}) &= \mathbf{A}; \phi(\mathbf{z}, \mathbf{B}) = \mathbf{A}; \end{aligned}$$

Now looking at the frequency of each letter we can find the expected frequency of **A** in the ciphertext. Using the equation above we will have the following:

$$\text{Probability}(\text{ciphertext } \mathbf{A}) =$$



$$\begin{aligned}
& \underbrace{k}_{(.007)} \underbrace{A}_{(.08)} + \underbrace{r}_{(.064)} \underbrace{S}_{(.066)} + \underbrace{y}_{(.02)} \underbrace{O}_{(.074)} + \underbrace{p}_{(.023)} \underbrace{T}_{(.084)} + \underbrace{t}_{(.084)} \underbrace{P}_{(.023)} + \\
& \underbrace{o}_{(.074)} \underbrace{Y}_{(.02)} + \underbrace{s}_{(.066)} \underbrace{R}_{(.064)} + \underbrace{a}_{(.08)} \underbrace{K}_{(.007)} + \underbrace{b}_{(.016)} \underbrace{Z}_{(.002)} + \underbrace{c}_{(.034)} \underbrace{X}_{(.002)} + \\
& \underbrace{d}_{(.038)} \underbrace{W}_{(.016)} + \underbrace{e}_{(.121)} \underbrace{V}_{(.01)} + \underbrace{f}_{(.02)} \underbrace{U}_{(.031)} + \underbrace{g}_{(.021)} \underbrace{Q}_{(.001)} + \underbrace{h}_{(.046)} \underbrace{N}_{(.074)} + \\
& \underbrace{i}_{(.078)} \underbrace{M}_{(.027)} + \underbrace{j}_{(.002)} \underbrace{L}_{(.045)} + \underbrace{l}_{(.045)} \underbrace{J}_{(.002)} + \underbrace{m}_{(.027)} \underbrace{I}_{(.078)} + \underbrace{n}_{(.074)} \underbrace{H}_{(.046)} + \\
& \underbrace{q}_{(.001)} \underbrace{G}_{(.021)} + \underbrace{u}_{(.031)} \underbrace{F}_{(.02)} + \underbrace{v}_{(.01)} \underbrace{E}_{(.121)} + \underbrace{w}_{(.016)} \underbrace{D}_{(.038)} + \underbrace{x}_{(.002)} \underbrace{C}_{(.034)} + \\
& \underbrace{z}_{(.002)} \underbrace{B}_{(.016)} \approx .03271.
\end{aligned}$$

Now we multiply this number by 97 (since there are 97 characters in K4) and we have our expected frequency of **A**  $\approx 3.17287$ .

Repeating this process for each letter we are able to calculate a Chi-Square value of  $\chi^2 \approx 22.72929$ . Taking a look at a Chi-Square distribution table with 25 degrees of freedom and reading along the row we find our value of 22.72929 to be between .900 (16.4734) and .100 (34.3816) [17, p. 850, 851]. This means that 90% of the time we can expect a value higher than 16.4734 and 10% of the time we can expect a value higher than 34.3816. Thus we cannot rule out the possibility that the fourth part was encrypted with a running key of transposed English.

## CHAPTER 6

### CONCLUSION

This “intelligence gathering” has proved itself a mystery. Many minds have brainstormed possibilities on how to fully decrypt the sculpture. Jim Sanborn may decide to reveal further hints in the future, but for now we can only work with what has been provided.

Along with discussing the first through third sections of the sculpture, we have discussed the shift cipher, Vigenère cipher, and transposition cipher. We have shown how the fourth section is not likely to be a Vigenère cipher or a Vigenère cipher with key ***ABCD*** followed by a transposition. We also completed a Chi Square statistic on a running key of transposed English that revealed some likeliness as a possible method of encryption of the fourth section.

Hopefully a conclusion will be reached soon and the mystery will be revealed. According to Sanborn, “*It doesn’t mean you’re going to understand it or it will be completely laid out before you... It will not be plain as day, ever* [20].”

## BIBLIOGRAPHY

- [1] Bardhaven. *The Nuance of Iqlusion*, 19 Jan. 2012,  
<http://bardhaven.wordpress.com/2007/08/14/the-nuance-of-iqlusion/>
- [2] Devore, Jay L. *Probability and Statistics for Engineering and the Sciences*, Third Edition, Brooks/Cole, Belmont, CA, 1991.
- [3] Dunin, Elonka. *Artist/Sculptor Jim Sanborn* 23 Feb. 2012,  
<http://elonka.com/kryptos/sanborn.html>
- [4] Dunin, Elonka. *Frequently Asked Questions About Kryptos*, 23 Feb. 2012,  
<http://elonka.com/kryptos/faq.html>
- [5] Dunin, Elonka. *Transcript of Kryptos Sculpture*, 23 Feb. 2012,  
<http://elonka.com/kryptos/transcript.html>
- [6] Eck, David and Ryan, Jim. *The Chi Square Statistic*, Math Beans Project. 19 Jan. 2012, <http://math.hws.edu/javamath/ryan/ChiSquare.html>
- [7] Gaines, Helen Fouché. *Cryptanalysis: a study of ciphers and their solution*, Dover Publications, Inc., New York, 1956.
- [8] Garrett, Paul. *Making and Breaking Codes: An Introduction to Cryptology*, Prentice Hall, Upper Saddle River, NJ, 2001.
- [9] G. & C. Meriam Company. *New Collegiate Dictionary*, A Merriam-Webster, Springfield, Massachusetts, 1980.
- [10] Gillogly, Jim. Kryptos Pix. Kryptos Sculpture, Pictures copyright (used with permission, from 12 Feb. 2012), 1999. <http://www.voyrich.net/Kryptos/>
- [11] Kahn, David. *The Code Breakers*. Scribner, New York, 1969.
- [12] Melvin, Joy. *Kryptos-CTV 2005*, Interview with Jim Sanborn, 3 Mar. 2012,  
[http://www.youtube.com/watch?v=ayXPI73OJ\\_o&feature=related](http://www.youtube.com/watch?v=ayXPI73OJ_o&feature=related)
- [13] Monico, Chris. Personal communication, 4 Jan. 2012.
- [14] Sanborn, Jim. *Jim Sanborn: The Artist's Official Site*, 23 Feb. 2012,  
<http://jimsanborn.net/>

- [15] Schwartz, John. *Clues to Stubborn Secret in C.I.A.'s Backyard*, appeared in print on November 21, 2010, on page A1 of the New York edition,  
<http://www.nytimes.com/2010/11/21/us/21code.html>
- [16] Schwartz, John. *Cracking the Code of a CIA Sculpture*, appeared in print Monday, July 19, 1999; Page A01,  
<http://www.washingtonpost.com/wp-srv/national/daily/july99/kryptos19.htm>
- [17] Wackerly, Mendenhall, Scheaffer. *Mathematical Statistics with Applications*, 7th Edition, Brooks/Cole Cengage Learning, Belmont, CA, 2008.
- [18] Waterloo Maple (Maplesoft), Computer algebra system with proprietary commercial software, version 15.
- [19] Wilson, John B. *Kryptos: The Sanborn Sculpture at CIA Headquarters- John's Collected Kryptos Hints*. 1999 NPR-All Things Considered interview. 5 Feb. 2012, <http://austininc.com/SciRealm/KryptosHints.html>
- [20] Zetter, Kim. *Kryptos Artist to Reveal Rare Clue to Baffling CIA Sculpture*, 23 Feb. 2012, <http://www.wired.com/threatlevel/2010/11/kryptos-clue/>
- [21] Zetter, Kim. *Questions for Kryptos' Creator*, 19 Jan. 2012,  
<http://www.wired.com/techbiz/media/news/2005/01/66333>
- [22] Zetter, Kim. *Solving the Enigma of Kryptos*, 15. Mar 2012,  
<http://www.wired.com/culture/lifestyle/news/2005/01/66334>
- [23] Zetter, Kim. *Typo Confounds Kryptos Sleuths*, 23 Feb. 2012,  
<http://www.wired.com/science/discoveries/news/2006/04/70701>