LING 106: Homework 3 Answer

1. CRYPTOGRAPHY

1.1. Deciphering Texts

Language is ergodic, so we expect the letter frequencies in a 365-letter text to approximate those in a longer text. We'd expect 46 E's (12.5% of 365 = 45.7), 34 T's (9.26% of 365 = 33.8), and so on.

Expected	Text (a)	Text (b)	Text (c)
46: E	27: GV	50: E	45: Y
34: T	25: T	33: N	33: Q
29: A	24: F	31: T	32: E
28: O	23: H	27: AI	26: J
27: I	21: R	25: O	25: I
26: N	17: C	24: S	23: R
24: S	16: I	20: D	22: A
22: R	15: X	18: H	21: G
20: H	14: UY	17: R	20: M
15: L	13: EL	14: L	19: BO
14: D	12: BNP	12: C	16: C
11: C	11: J	9: BY	11: PV
10: U	10: AM	8: FUW	9: S
9: M	9: KQ	7: G	7: T
8: F	8: W	6: P	6: X
7: PGW	7: DZ	5: MV	5: H
6: BY	6: S	1: JX	4: KU
4: V	3: O		3: N
2: K			2: F
1: JX			1: LW

- Text (b) has roughly the same frequencies as the expected English frequencies, so it's probably a rearrangement cipher, which would preserve the exact letter counts.
- Text (c) has the right percentages, but the actual letters are scrambled, so it's probably a monoalphabetic cipher, which preserves relative percentages (and Y, being most common, is probably E; Q is probably T; and so on).
- The frequencies in text (a) are far too skewed—the most common letter appears half as much as we'd expect the most common English letter to appear, there's no significant gap in the top four or five frequencies, and so on. This is presumably a polyalphabetic substitution cipher.

For reference, the texts are all from *The Da Vinci Code*

a. "You are describing only a small portion of the Opus Dei population," Aringarosa said. "There are many levels of involvement. Thousands of Opus Dei members are married, have families, and do God\'s Work in their own communities. Others choose lives of asceticism within our cloistered residence halls. These choices are personal, but everyone in Opus Dei shares the goal of bettering the world by doing the Work of God. Surely this is an admirable quest."

   Encryption method: Vigenère cipher, keyword ENCRYPT

b. Langdon was feeling anything but fortunate, and coincidence was a concept he did not entirely trust. As someone who had spent his life exploring the hidden interconnectivity of disparate emblems and ideologies, Langdon viewed the world as a web of profoundly intertwined histories and events. The connections may be invisible, he often preached to his symbology classes at Harvard, but they are always there, buried just beneath the surface.

   Encryption method: Take every fifth letter starting with the first, then the second, etc.

c. The exact length, if Langdon recalled correctly, was around fifteen hundred feet, the length of three Washington Monuments laid end to end. Equally breathtaking was the corridor's width, which easily could have accommodated side by side passenger trains. The center of the hallway was dotted by the occasional statue or colossal porcelain urn, which served as a tasteful divider and kept the flow of traffic moving down one wall and up the other.

   Encryption method: Simple substitution cipher, keyword ENCRYPT

1.2. Enciphering Texts

Week 1
(i) For each word, alphabetize the letters in the word.
(ii) Response: `egiv em eht eky adn ill do eht bjo`

Partial credit for recognizing that each word was rearranged, even if you thought it was random. But I'm a little suspicious of anyone who said that the rearrangement was random, and yet rearranged **the** into `eht` both times. Not very random, is that?

Week 2
(i) Reverse each pair of letters.
(ii) Response: `igev em htk eye nai dll od htj ebo`

Week 3
(i) Caesar shift back three (or forward 23).
(ii) Response: `dfsb jb qeb hbv xka fii al qeb gly`

Week 4
(i) Vowels: **a** = 5, **e** = 4, **i** = 3, **o** = 2, **u** = 1
Consonants: **b** = z, **c** = y, ..., **z** = b (i.e. replace the *n*th consonant in the alphabet with the *n*th-to-last consonant)
(ii) Response: `v3g4 p4 ht4 r4c 5nx 3qq x2 ht4 s2z`

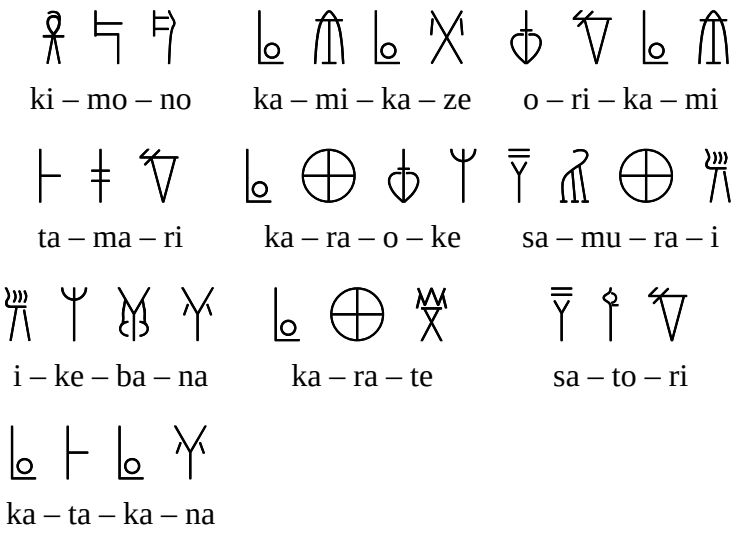
This was one where finding the pattern was particularly important; recognizing that it was a simple substitution cipher wouldn't help you encipher *give* or *job*.

Week 5
(i) For each word, reverse the word and then Caesar shift forward 10
(ii) Response: `ofsq ow ord iou nxk vvs yn ord lyt`

This was hard; the first time I saw it, it took me a little time to determine what was going on here.

2. LINEAR B

The words, annotated:



Thus: **Naomi** =  and **mimosa** = 

And finally,  = *ma-i-ze* = “maize”, except that the word “maize” isn’t pronounced “ma, i, ze”—these represent spoken syllables, not strings of letters.