

Student's Name:

Course Name: Linguistics 484

Set 16 Feb 2008, Due in class W 23 Feb 2008

Teacher's Name: Chris Brew

---

**Solve the following keyword based Vigenere cipher. (The ciphertext is on website as homework3.txt)**

DQVZE CDUPI SWLPP GMSJI CTMYS DPOVR NMWZH ZFUJT SLLPT CEMWE  
LIHVJ MODRJ WICXL FXIMJ SBHMC CEGKB SZGRL IWJKL WWQXE KSTZB  
YVKVZ RXYSB XTIPW ANVRY SZOGC UWKTK LVJES VXZOZ TOIZS UAMIG  
SZTIK KGLTZ NLKBL AINDM LBYVK NCJQK ZMSZB RYWYD RNZQN CMNSA  
MJVEG VNZMR ELPVH KZMCZ AZDTM ZRFDI NFSWH QZPWY SVONX FLPCJ  
ADWWG ZVKZM CVMCA VRHCX JIGZA ZDTMZ XYWNL OLFET PNWRA ZXTFF  
VGHDP CKQYF IEVTP NWCQI YYHVU IJVUR VQDNS CNMTI XYWET IHXWV  
PMEKW LMTXY WNLGP NZQNC MJAVQ DRZLM TNEPL PLOXY WTTWV RJGTN  
YEWVO DRXLP PDHVS TTNXJ SZRPI KZIEO LVZMI VKFFI WMSFE ALMIR  
FMNZW JSZJA VFEWQ VFJGT FOIJH INZSI SBWZE JLWQJ YIAVE PMKAW  
YJJJH INZXY WGCZE JGVEC EKSBC DEEYC WVVFJ XPIXR YWYVP IGWXD  
WZFKZ IGVAD LWPVL PPHCJ LQNGG CSQXO LRLBS ZMIWK DOEJQ ZPQIR  
DAEJX YWULX MIUCW VVTZI XWIIU WYOEZ FOYBE XJMLO GZJKF GEITW  
ZFAYG APNTZ FMTNG FFBTI YFMAL IHNZQ NCJFD TZRWK ZMNJQ GDMEZ  
GZJKW ZSWLP PRECD AMPXK ZMTMX VKBTH SEQQD NYJHM NOXYW QCRSI  
VAZWW TMZPO LZKKJ XPZUI WWSFC QDBSU DMEDX JMNQD GVFWH ASIEM  
EJVVH MLOXY WKVWV JAKOD GKMUE CICAJ CVVPA ALNTY WZPRL FKMP  
ETLKP IXVJQ DVRPG VPJJZ LASZB RYWYN EEVES JWVUQ CXYDX MCZRT  
WQDDR RUKPN WZTTP

## Frequency analysis: compute the index of coincidence using these frequencies (needs a calculator or similar)

Letter frequencies in the message are

A:27 B:15 C:34 D:33 E:40 F:29 G:27 H:18 I:45 J:44 K:34 L:40 M:47 N:38  
O:19 P:40 Q:27 R:29 S:35 T:38 U:14 V:52 W:59 X:34 Y:34 Z:63

First add the frequencies. There are a total of 915 characters, so your total should be 915.

Now, for each frequency  $f$ , calculate the quantity  $f(f-1)$ . e.g. for 27, should be  $27(26) = 702$ .

Add these numbers together. Your total should be 35010

Divide this by  $914 * 915$ . Your total should be a small number. Where does it fit in the table below?

Number of alphabets	IC
1	0.066
2	0.052
5	0.044
10	0.041
large	0.038

This gives you a rough idea of the number of alphabets.

## Kasiski-Babbage analysis

Here are the positions of all four letter repeated sequences in the ciphertext. The first line means that AZDT happens two times, first at position 220, second at position 274.

From this information, using the ideas in the lecture and the Singh, work out what the period of encryption probably is, and hence how many alphabets there are.

AZDT [220, 274]

BRYW [184, 874]

BYVK [99, 171]

CWVV [529, 613]  
DTMZ [222, 276]  
GZJK [640, 700]  
HINZ [474, 504]  
JHIN [473, 503]  
KZMC [215, 257]  
LOXY [386, 806]  
NZQN [191, 365, 677]  
OXYW [387, 741, 807]  
QNCM [193, 367]  
RYWY [185, 539, 875]  
SWLP [10, 706]  
SZBR [182, 872]  
TPNW [289, 313]  
WLPP [11, 707]  
XYWN [280, 358]  
YWNL [281, 359]  
ZAZD [219, 273]  
ZBRY [183, 873]  
ZDTM [221, 275]  
ZQNC [192, 366, 678]  
ZXYW [279, 507]

**Finally, use the information you have about how many alphabets to group the message into columns and solve the code.**