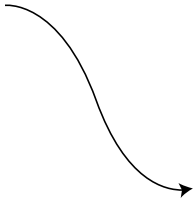
La lezione di oggi è tutta incentrata sulla seguente domanda: **come si può rompere una cifratura per sostituzione**?

Per rispondere alla domanda consideriamo un messaggio basato su una tavola esclusivamente alfabetica di 26 simboli e cifrato per sostituzione; mettiamoci nei panni di un ipotetico intercettatore, ignaro del testo in chiaro e dell'alfabeto cifrante ma a conoscenza della lingua del messaggio (l'italiano) e della generica tecnica crittografica usata: la sostituzione monoalfabetica. Per maggiore praticità espositiva considereremo soltanto un estratto composto dalle prime 180 lettere del testo (*vedi figura in basso*): si tratta di una scelta squisitamente didattica, ovviamente un vero crittoanalista non rinuncerebbe mai ad una parte del messaggio (che consta complessivamente di più di settantamila lettere!).

FRVSLSXYLRMCBCKKDQCNXAAKLFDLSD
KLDMDJDYYCSSDXEELFLDKCBLFCPAQC
MLKFDNRNQRTTLVRQLRBCKKRVSDSRTL
VSDKDBCKLACQDYLRMCBCKKDVVCPAKC
DFRVSLSXCMSCFGCMCKKDVCBXSDBCKB
LFCPAQCGDDNNQRTDSRKDFRVSLSXYLR
MCBCKKDQCNXAAKLFDLSDKLDMDTLVSD
KDWTLLLBLVNRVLYLRMCELMDKCBCKKD
FRVSLSXYLRMCNQRPXKJDKDFRVSLSXY
LRMCBCKKDQCNXAAKLFDLSDKLDMDMCK
VCJXCMSCSCVSRNQLMFLNERMBDPCMSD
KLDQSKLSDKLDCXMDQCNXAAKLFDBCPR
FQDSLFDERMBDSDVXKKDTRQRKDVRTQD
MLSDDNNDQSLCMCDKNRNRKRFGCKDCVC
QFLSDMCKKCERQPCCMCLKLPLSLBCKKD
FRVSLSXYLRMCDQSKDQCNXAAKLFDQLF
RMRVFCCJDQDMSLVFCLBLQLSSLLMTLR
KDALKLBCKKXRPRVLDFRPCVLMJRKRVL
DMCKKCERQPDYLRMLVRFLDKLRTCVLVT
RKJCKDVXDNCQVRMDKLSDCQLFGLCBCK
DBCPNLPCMSRBCLBRTCQLLMBCQRJDAL



FRVSLSXYLRMCBCKKDQCNXAAKLFDLSD
KLDMDJDYYCSSDXEELFLDKCBLFCPAQC
MLKFDNRNQRTTLVRQLRBCKKRVSDSRTL
VSDKDBCKLACQDYLRMCBCKKDVVCPAKC
DFRVSLSXCMSCFGCMCKKDVCBXSDBCKB
LFCPAQCGDDNNQRTDSRKDFRVSLSXYLR

7.1) Testo cifrato alla mano, come affrontereste questo problema di crittoanalisi?

Ricordando che la sostituzione scambia i simboli delle lettere ma non la loro posizione, possiamo dedurre che alcune caratteristiche ortografiche della lingua italiana si siano trasferite al testo cifrato. Vediamo in seguito un esempio che chiarisce il senso di questa considerazione.
In italiano non esistono parole con tre lettere uguali consecutive, ciononostante può accadere che tre lettere identiche si succedano all'interno di una frase: considera ad esempio "gli z**ii** **i**nvitati", "uno z**oo** **o**rrendo" o, rinunciando alla d eufonica, l'espressione "fatt**a** **a** **a**rte". Tra l'altro, l'assenza di punteggiatura fa sì che queste triplette possano crearsi anche tra due frasi separate da un punto ([…] hanno dichiarato le due r**ee.** **E**videntemente esse […]). Visto che in italiano queste eccezioni riguardano esclusivamente le vocali, possiamo ricavare da un'ipotetica tripletta QQQ di un testo cifrato i seguenti indizi:

- All'interno della tripletta vi è un'interruzione tra parole
- Il simbolo che si ripete nella tripletta (la Q) è una vocale

L'ultimo indizio è importantissimo e permette di ridurre drasticamente il numero di tentativi necessari a rompere la cifratura.

Il caso appena considerato purtroppo non si applica al testo che stiamo considerando, suggerisce però una linea d'azione legato alla seguente domanda: **nel codice cifrato quali tracce ha lasciato l'italiano**?

Iniziamo l'analisi contando i vari simboli presenti nella porzione di testo cifrato che stiamo considerando: sopra tutte si notano la C (22 occorrenze) e la D (21).

Occorrenze della C

FRVSLSXYLRMCBCKKDQCNXAAKLFDLSD
KLDMDJDYYCSSDXEELFLDKCBLFCPAQC
MLKFDNRNQRTTLVRQLRBCKKRVSDSRTL
VSDKDBCKLACQDYLRMCBCKKDVVCPAKC
DFRVSLSXCMSCFGCMCKKDVCBXSDBCKB
LFCPAQCGDDNNQRTDSRKDFRVSLSXYLR

Occorrenze della D

FRVSLSXYLRMCBCKKDQCNXAAKLFDLSD
KLDMDJDYYCSSDXEELFLDKCBLFCPAQC
MLKFDNRNQRTTLVRQLRBCKKRVSDSRTL
VSDKDBCKLACQDYLRMCBCKKDVVCPAKC
DFRVSLSXCMSCFGCMCKKDVCBXSDBCKB
LFCPAQCGDDNNQRTDSRKDFRVSLSXYLR

Dopo aver considerato le lettere singole, consideriamo le doppie. A fianco (→) sono messi in rilievo anche i simboli posti immediatamente prima e dopo (anche essi portatori di informazioni importanti).

Presenza delle doppie

FRVSLSXYLRMCBCKKDQCNXAAKLFDLSD
KLDMDJDYYCSSDXEELFLDKCBLFCPAQC
MLKFDNRNQRTTLVRQLRBCKKRVSDSRTL
VSDKDBCKLACQDYLRMCBCKKDVVCPAKC
DFRVSLSXCMSCFGCMCKKDVCBXSDBCKB
LFCPAQCGDDNNQRTDSRKDFRVSLSXYLR

Triplette composte da simboli identici non ve ne sono, salta però subito all'occhio la tripletta ricorrente CKK, quasi sempre preceduta da una B e/o seguita da una D. Come prima ipotesi verrebbe da dire che K è una consonante, e B, C, D tre vocali.

Ripetizione di gruppi di tre lettere

FRVSLSXYLRMCBCKKDQCNXAAKLFDLSD
KLDMDJDYYCSSDXEELFLDKCBLFCPAQC
MLKFDNRNQRTTLVRQLRBCKKRVSDSRTL
VSDKDBCKLACQDYLRMCBCKKDVVCPAKC
DFRVSLSXCMSCFGCMCKKDVCBXSDBCKB
LFCPAQCGDDNNQRTDSRKDFRVSLSXYLR

È difficile trarre informazioni più precise dai dati a disposizione se non conosciamo le caratteristiche ortografiche salienti della nostra lingua: quali sono le lettere più frequenti, quali sono le doppie e più in generale le coppie di lettere più frequenti, i gruppi di tre lettere ricorrenti e così via?

Per rispondere a questa domanda è opportuno considerare un testo di riferimento possibilmente lungo e che possa replicare a grandi linee il lessico e il contesto del testo da decifrare. Se per esempio dobbiamo decifrare un testo che presumibilmente conterrà molte abbreviazioni, il numero di vocali potrebbe essere inferiore al normale e sarà opportuno confrontare il codice con un testo analogo. In assenza di informazioni di questo tipo

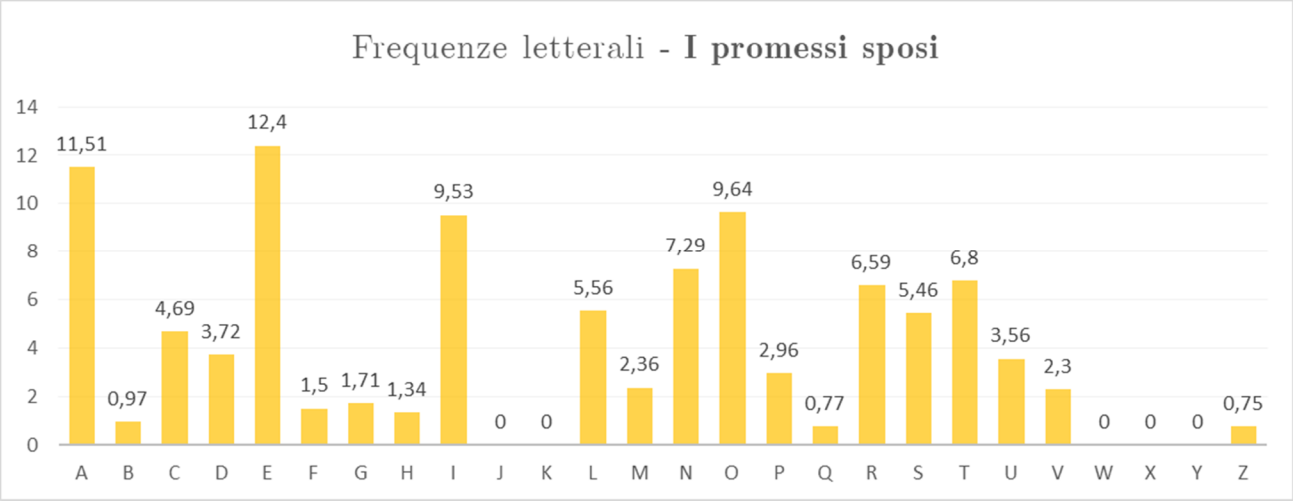
bisogna usare un testo qualsiasi, della lingua della cifra. Consideriamo come prototipo il testo di riferimento dell'italiano moderno, cioè *I promessi sposi* di Manzoni.

Eseguiamo ora un'analisi delle frequenze, stabiliamo cioè la quota percentuale di ciascuna lettera (con considerando accenti oi dieresi) nel testo privato dei segni di interpunzione:

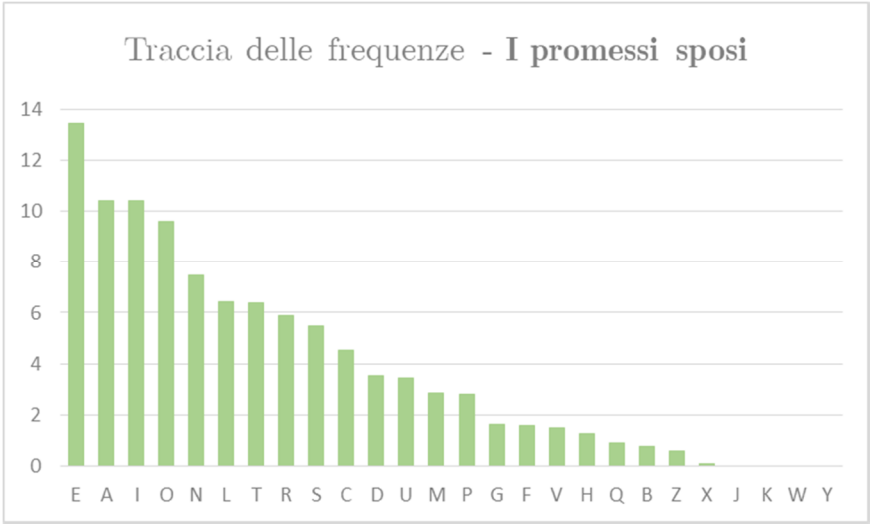
Frequenze letterali ne *I PROMESSI SPOSI*

A	11,51	J	0	S	5,46
B	0,97	K	0	T	6,8
C	4,69	L	5,56	U	3,56
D	3,72	M	2,36	V	2,3
E	12,4	N	7,29	W	0
F	1,5	O	9,64	X	0
G	1,71	P	2,96	Y	0
H	1,34	Q	0,77	Z	0,75
I	9,53	R	6,59		

Le statistiche riportate sopra offrono un ottimo spaccato, dal momento che il romanzo di Manzoni è composto da più di un milione di lettere. L'analisi diventa più veloce se si rappresentano i dati in forma di istogramma:



Un altro grafico utile per la decrittazione è la cosiddetta "Traccia della frequenze", che riporta gli stessi dati di prima ma in ordine decrescente.



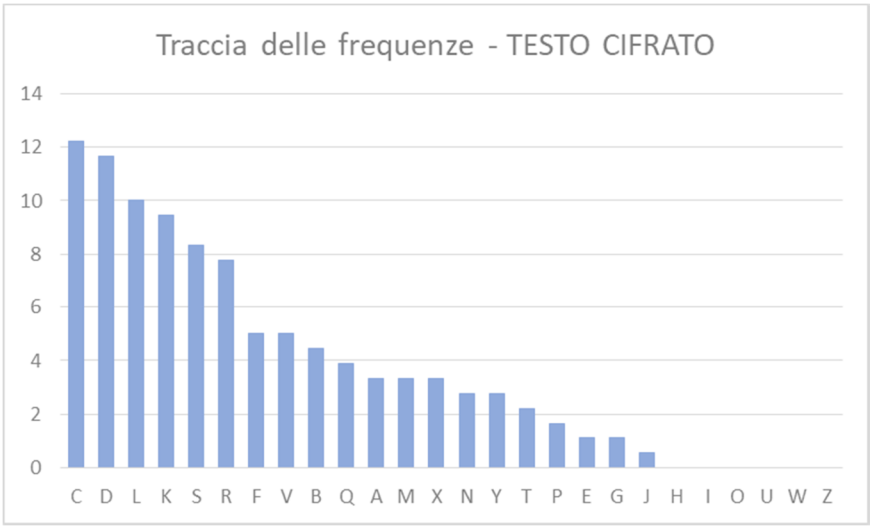
Vediamo altre due statistiche tratte di *I promessi sposi* relative rispettivamente alle coppie di lettere (dette **bigrammi**) più ricorrenti e alle doppie:

Bigrammi e doppie ne *I PROMESSI SPOSI*

Bigramma	Percentuale
ER	1,91%
ON	1,68%
EN	1,54%
AN	1,52%
RE	1,44%
TO	1,38%
ES	1,36%
DI	1,31%
AR	1,37%

Doppia	Percentuale
LL	0.86%
TT	0.73%
SS	0.64%
CC	0.38%
AA	0.3%
EE	0.3%
BB	0.19%
RR	0.18%
NN	0.18%

Confrontando ora con le informazioni sul testo cifrato e sulla traccia delle frequenze riportata in basso, possiamo iniziare a fare le prime ipotesi.

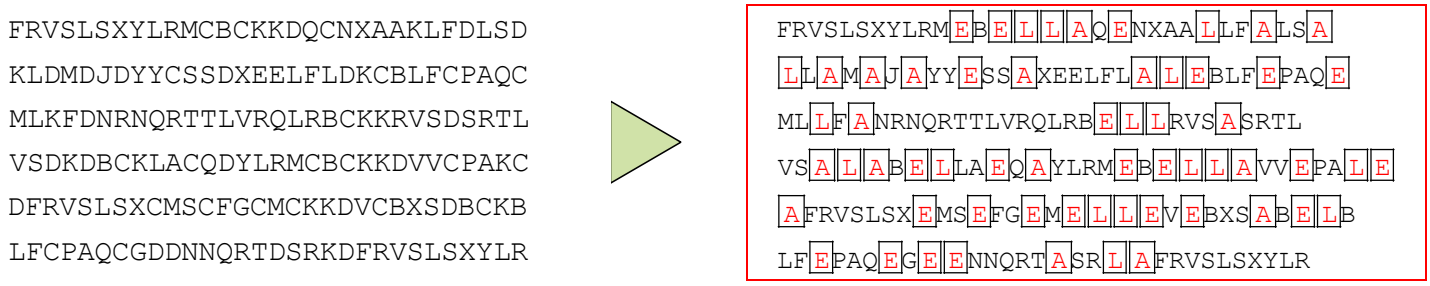


L'idea più ovvia, anche se non suffragata da una robusta evidenza statistica, è $C \to E$, $D \to A$, seguita a ruota da $C \to A$, $D \to E$. Per velocizzare l'analisi imbocchiamo da subito la strada corretta: si tratta della seconda scelta, che completiamo con l'ipotesi $K \to L$ (ricordo che L è la consonante più frequente nel nostro testo di riferimento e che abbiamo precedentemente ipotizzato che K rappresentasse proprio una consonante).

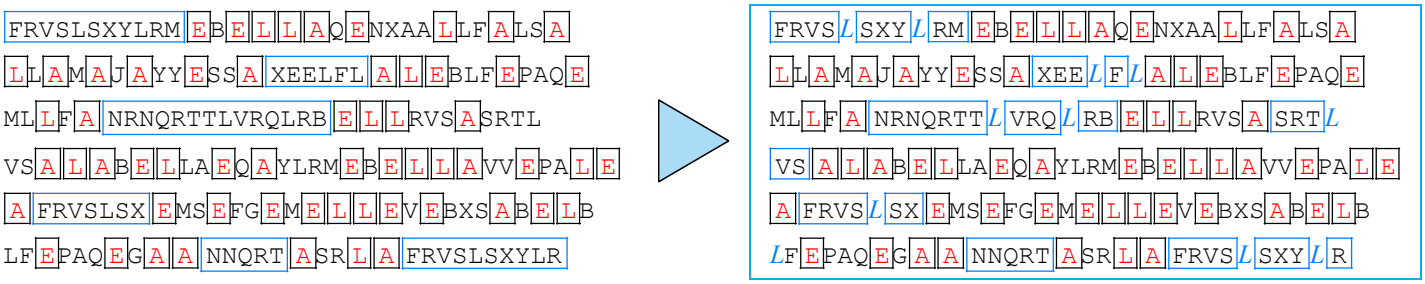
TABELLA DEL DECIFRATORE (o *TABELLA DI DECIFRAZIONE*)

Alfabeto CIFRATO	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
Alfabeto CHIARO			**A**	**E**							**L**															

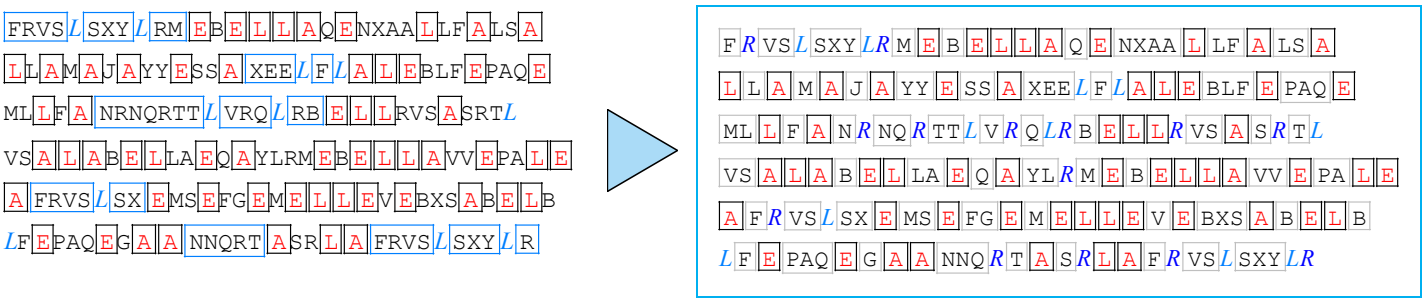
Operando quindi le sostituzioni indicate si arriva a…



Se le nostre ipotesi fin qui sono corrette, dovrebbe essere possibile individuare le vocali mancanti, la *I*, la *O* e la U. Nello schema in basso a sinistra sono messe in rilievo (nei riquadri azzurri) le sequenze prive di A ed E (sempre che le nostre ipotesi fino qui siano giuste). Una parola in italiano non può contenere più di 4 consonanti di fila, per cui dobbiamo riuscire a "spezzare" queste sequenze con qualche vocale. La L sembra la miglior candidata:



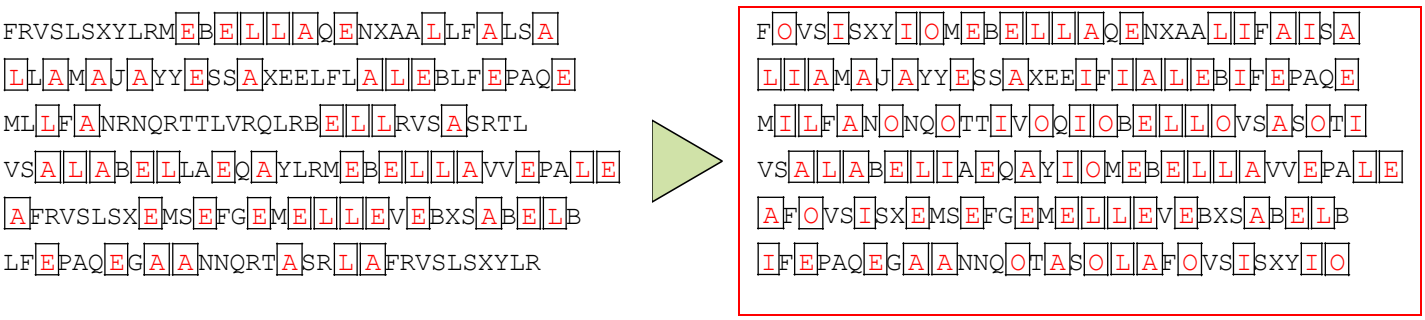
Restano le due sequenze "lunghe" NRNQRTT e NNQRT, che suggeriscono come possibili vocali la N, la Q e la R. La traccia delle frequenze riportata in alto dà alla R maggiore probabilità.



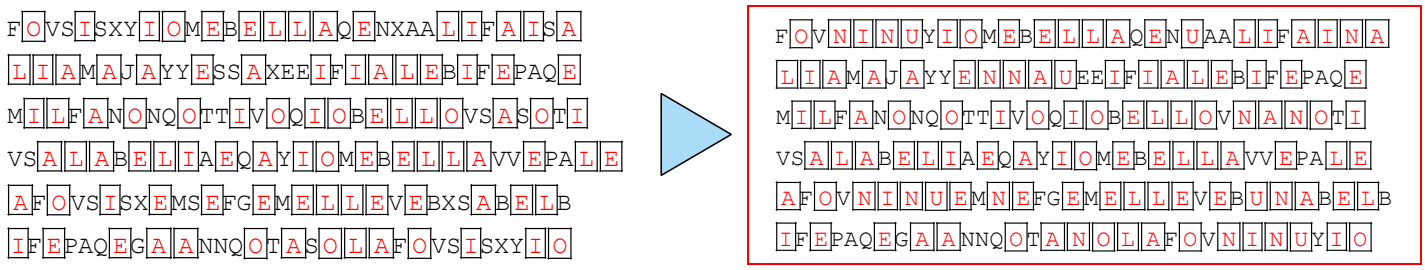
Siamo ora riusciti a spezzare quasi tutte le catene, il che depone a nostro favore. L'unica "quadrupla" superstite NXAA potrebbe suggerire una U al posto della N o più probabilmente al posto della X. Osservano bene le L e le R, si vede che il bigramma LR occorre 3 volte, mentre RL mai. Ciò fa pensare che LR rappresenti IO (ben più frequente in italiano di OI). Vediamo dove ci porta questa speculazione:

TABELLA DEL DECIFRATORE (o *TABELLA DI DECIFRAZIONE*)

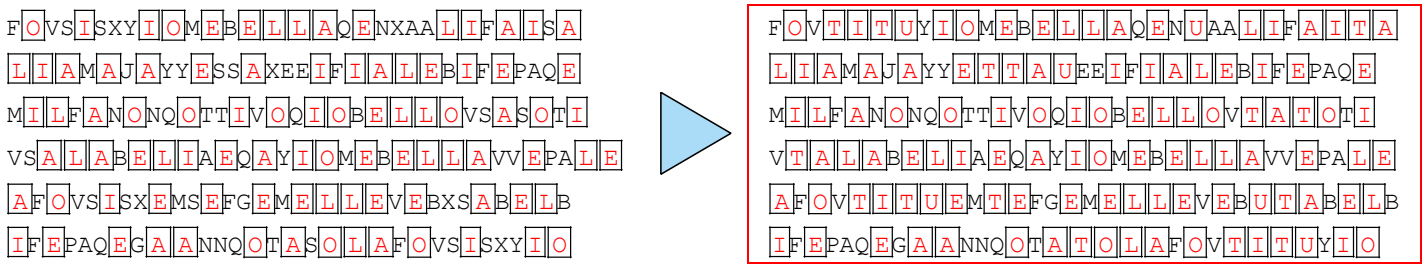
Alfabeto CIFRATO	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Alfabeto CHIARO			**A**	**E**							**L**	**I**						**O**								



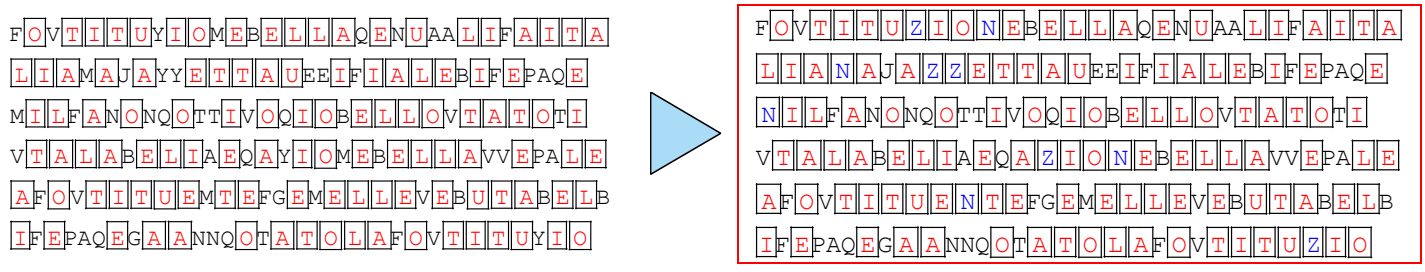
Il messaggio non è ancora decrittabile, è tempo di trovare le consonanti. Delle lettere non ancora assegnate, la S ha la maggior frequenza. Proviamo a sostituirla con la N e poi a perseguire la "vecchia ipotesi" X→U.



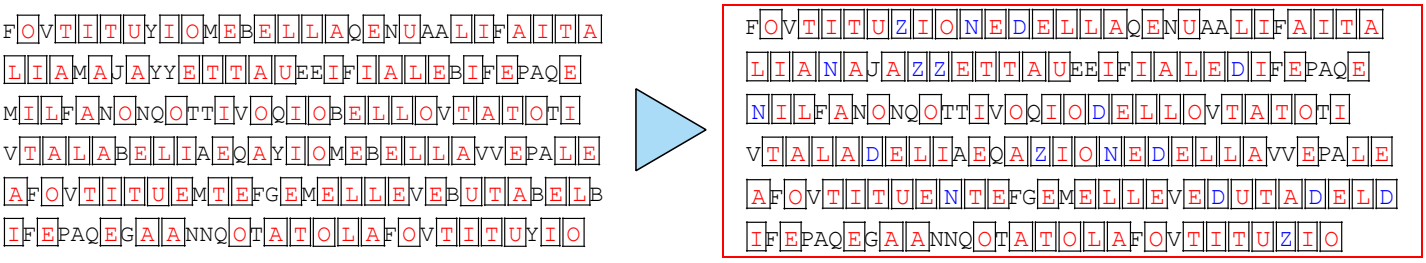
Chiaramente qualcosa non sta funzionano: NINU, ENNAU, NINUE non sembrano essere corretti. Facciamo quindi un passo indietro e proviamo S→T, X→U:



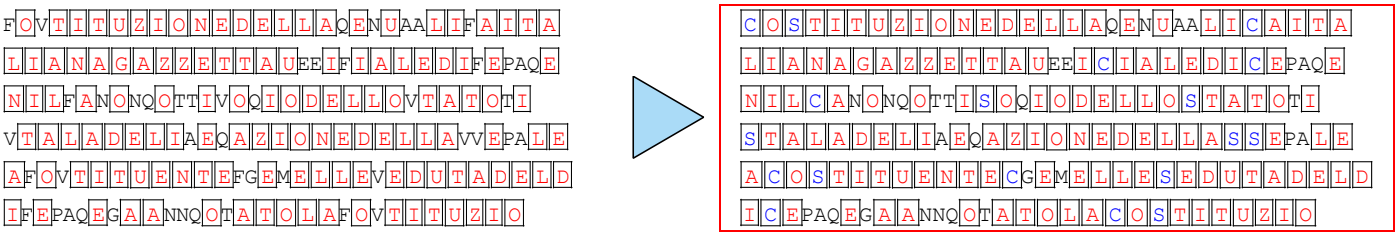
La stringa YIOME ripetuta due volte (e ripresa alla fine) potrebbe essere il finale ZIONE, il che ci porterebbe a …

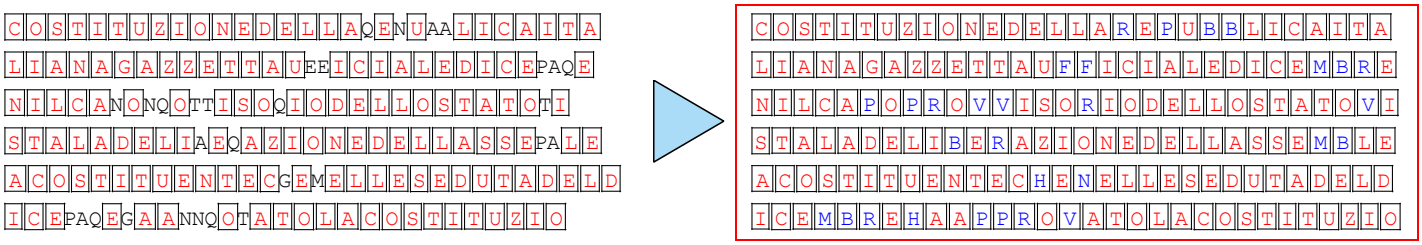


A questo punto BELLA sembra dover portare a DELLA …



…e J A Z Z E T T A suggerisce G A Z Z E T T A. Come si vede, siamo arrivati ad una fase in cui le lettere cadono una ad una.





Il messaggio risultante, non privo di qualche errore (dovuto “alla perdita” dei numeri in fase di traduzione), è il cappello introduttivo della pubblicazione in Gazzetta Ufficiale della Costituzione della Repubblica Italiana. Abbiamo così decrittato, non senza fatica e con qualche trucco, il nostro primo messaggio cifrato per sostituzione. Vediamo come ve la cavate con i prossimi:

Problema 7.1) Da quale libro è stato estratto questo testo?

MESSAGGIO INTERCETTATO 1 (testo cifrato con una sostituzione monoalfabetica)

```
HQJWGNLQHMWJGWEFCQHPDPWIPGAGDWINGJWFCQOQKNSWZZQLPDWINQHMWSPGNIWOGWCPL
WOGHQJWKNVGJVPNQHWCGKNGOQDFGKNCWLPCGNNQEKVPFQPJPWPSPHHQDPDWFFQCPNQCGI
GZZPGOWFWEVGIDPGFQHWCGKNGOQDFGKNSWZZQLPDWINQNQNWCGKNDWINQLPDKEEQJGKNE
WJSDPHWSWZZQLGHGFGEFGLPYKWDDPHMWLPNOWCNQEPJWFFQNQNWDDWEFKAWWNWPHGJPNW
FFPSWCGHHWNLWCWPDAKQHQWSWCCPEHGDLGCWDWEFGNZWNQNEQHQJWGNLGEEWJGPDAGFFQ
IDPWHMWKNVWDIPQCNQYKWEFQSWZZQLPDWINQHGSPFQNWDDGVQFFWIGLPKNOWHHMPQAGDW
INGJWPDYKGDWGOWOGNQJWJGEFCGNFQNPQEWNQNHMWFKFFPDQHMPGJGOGNQJGWEFCQHPDP
WIPGSWCOPGLWDDGSKNFGLWDEKQNGEQHMWWCGEWJSCWDKEFCGWSGQNGZZGHQJWKNGHPDPW
IPGJGFKCGPDDKEFCGZPQNWEWNFPKNGOQHPNGEQFFPDWEQFFPDWGSSWNGJGWEFCQHPDPWI
PGWVVWOPEFQYKWDSWZZQLPDWINQEPCGDDWICQFKFFQWLGNLQEPKNGACWIGFPNGLPJGNPS
WCDGHQNFWNFWZZGVQCVQFFQGJWZZGOQHWYKWEFQDWINQWHGSPFGFQGFWJSQOQIDPQEWCO
PCJWNWSWCAGCWKNGIGJVGLPFGOQDPNQLWFFQAGFFQSCWEWEKVPFQDGEHPGGCCQFGFGSWC
HQJPNHPGCWGDWOGCIDPDGEHQCZGWGLPICQEEGCDQJGYKGNLQAKDPSWCDGEHPGCWGNLGCW
DGSCPJGGEHPGFGCPJGEWHQDVCGHHPQEQESWEQPNGCPGSWCHMWEWNFPKNGOQHPNGEQFFPD
WEQFFPDWHMWLPEEWCGHHQJGNLGNLQEPNQNJPSPHHMPGCFGNFQAQCFWAPIKCGFWOPHQJWC
PJGEWYKWDVKQNOWHHMPQLPJGWEFCQHPDPWIPGIPCQIDPQHHMPEJGCCPFPPNFQCNQGDDGE
FGNZGSWCOWLWCWLPLQOWJGPSQFWOGWEEWCWKEHPFGYKWDDGOQHPNGWNQNOPLWNWEEKNQI
KGCLQEQFFQPDVGNHQWNWEEKNQIKGCLQLWNFCQKNGCJGLPQHMWEFGOGEWJSCWHMPKEQWNW
EEKNQIKGCLQNWDHQCVWDDQLWPFCKHPQDPWLWDDGEWIGFKCGWNWEEKNQGSCPDKEHPQLPVQ
FFWIGSWCLGCWKNQHHMPGFGGNHMWEKDDGEFCGLGWNWEEKNQQLKNYKWMQHGSPFQLPEEWGDD
QCGCPLWNLQWICGFFGNLQEPDGSGCCKHHGEPOWLWHMWYKWDDGOQHPNGJWDGEQNAPIKCGFGP
QCPJWFFPGJQHPGDGOQCGCWWCPSCWEGDGEHPGPNJGNQFPCQIPKKNEQDWNNPEEPJQHQDSQE
KDSWZZQLPDWINQQMPFKJMGPAGFFQJGDWICPLQCGJJGCPHGNLQEPDGEQDPFGOQHPNGYKWE
FGOQDFGJGWEFCQHPDPWIPGCWEFQLPEFKHHQHQIDPQHHMPAKQCPLWDHGSQSWCDGSGKCGHQ
```

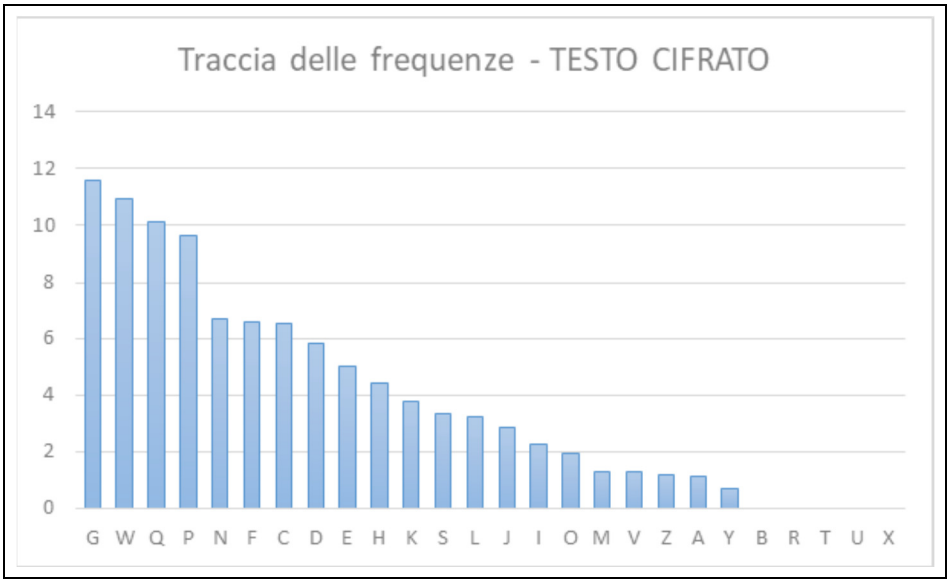
DDGVQHHGESGDGNHGFGWHQDDGDPNIKGIPKHPQNLQDQNPAPNQGDJWNFQHQJWKNJGEHMWCQN
WLGAQNFGNGGSSWNGCPWVVWDKEQLWDDGSGCQDGHQJPNHPQGLPCWFCWJGNLQWVGDVWFFGNL
QLGDDQESGOWNFQJGLPLQOWEGCGKEHPFGYKWEFGOQHPNGHMWMGLWFFQQMPWSSKCWYKPNQN
HWGNPJGOPOGHMWEPGSWCHGEQYKWEFQSWZZQLPDWINQHMWGVVPGPJSGCGFQGSPGNIWCWWG
DGJWNFGCEPHQJWKNVGJVPNQPQNQNDQSQEEQHCWLWCWYKWEFQDWINQWHHQDQYKPWKNSWZZ
QLPDWINQLGHGJPNWFFQHQJWFKFFPIDPGDFCPWGVKFFGCDQEKDAKQHQHWLGAGCVQDDPCWK
NGSWNFQDGLPAGIPQDPQLKNYKWHMWHPEPGNGEHQEFQLWNFCQYKGDHKNQEWHWNGEHQEFQYK
GDHKNQFGNFQSWIIPQSWCDKPQCGDGHHQJQLQPQWHQEPLPHWNLQGIIKGNFQHQNFKFFWLKWD
WJGNPYKWDSQOWCQSWZZQLPDWINQWEPSQEWGEVGFGHHMPGCDQEWNZGHGCPFGHQNFCQDWSG
CWFPLWDDGEFGNZGSQPEPJWEEWPNGEHQDFQSWCEWNFPCWEWHWCGYKGDHMWOQHPNGHMWEPD
GJWNFGEEWGESWFFQLKWJPNKFPWNKDDGHPNYKWJPNKFPWNKDDGLPWHPJPNKFPWNKDDGMQH
GSPFQLPEEWGDDQCGEAQCZGNLQEPLPCPLWCWWGCCKAAGNLQEPDGSGCCKHHGEPOWLWHMWYK
WDDGOQHPNGHMWMGLWFFQQMPJWDGEQNAPIKCGFGPQCPJWFFPGJQHPGDGOQCGCWWSWCHMWI
DPWCGWNFCGFQGLLQEEQKNGICGNSGKCGEPSCQOQGHGNFWCWDDGCWSWCAGCEPKNSQLPHQCG
IIPQPNFGNFQSQEGFGLGKNGSGCFWDGEHPGSCWEWPNJGNQDGSPGDDGSWCSPGDDGCWWFPCGC
WGSKDPJWNFQPDSWZZQLPDWINQJGNWDJWNFCWHMWDQSPGDDGOGPNEKWPNIPKEWNFPDGEQD
PFGOQHPNGHMWIDPLPEEWCPLWNLQEJWFFPFKJPAGPPDSPZZPHQCPNQEKDHQCSQYKWEFGOQ
DFGPDSQOWCQJGWEFCQHPDPWIPGHGLLWIPKHQJWAKDJPNGFQYKGNLQCPGSCPIDPQHHMPEP
FCQOQEWLKFQSWCFWCCGPDEKQOPEQSGCWOGFCGEAPIKCPFQWSWCAPNQDGSKNFGLWDNGEQL
PSGQNGZZGHQJWWCGYKGEPEWJSCWIDPWCGLPOWNFGFGFKCHMPNGLGDDGICGNSGKCGPPJGW
EFCQHPDPWIPGCWIGDGPDSWZZQLPDWINQGDEKQGJPHQIWSSWFFQPDYKGDWDQSCWNLWSWCA
GVVCPHGCEPKNVKCGFFPNQJGCGOPIDPQEQHMWEGSSPGVGDDGCWFPCGCLPEHMWCJGWAGCWP
EGDFPJQCFGDPPNYKWDSKNFQAKVKEEGFQGDDGSQCFGSGEEGFWSKCWLPEEWPDAGDWINGJWE
WNZGGOWCDGAQCZGLPCPZZGCEPPNSPWLPPDDKEFCGZPQNWKNOWHHMPWFFQFKFFQGCZPDDQ
PDYKGDWGOWOGNQJWIWSSWFFQGDDQCGWNFCQPNVQFFWIGKNOWHHMPWFFQFKFFQGCZPDDQP
DYKGDWGOWOGNQJWIWSSWFFQJGPCGIGZZPLWDOPHPNGFQYKGNLQDQOQDWOGNQAGCJQNFGC
WEKFKFFWDWAKCPWDQHMPGJGOGNQHQDEQSCGNNQJWLPSQDWNLPNGGJQFPOQLWDDGEKGSGC
CKHHGIPGDDGHMWEQJPIDPGOGJQDFPEEPJQGDDGSQDWNLPNGLPICGNFKCHQIWSSWFFQWCG
VPZZQEPEEPJQIKGPGHMPGJGCDQSQDWNLPNGLPOWNFGOGEKVPFQKNGVWEFPGWNQNHWCGSP
KOWCEQLPFWNWCDQVKQNIPQCNQJGEFCGNFQNPQLPEEWIWSSWFFQHMWHQEGAGFWHQEFPSWC
FWCCGPNEWINQDGVVGHQGDDWAQCJPHQDWVKQNSCQOPAGHHPGHMPOPMGSQCFGFQLGJWHQJS
GCIWSSWFFQDWIGJVWEGSSPGFWJGEFCGNFQNPQHMWEQNOWNKFQLGOQPSWCHMPWLWCOPKNA
GOQCWWHHQJPYKPSCQNFQGEWCOPCOPCWSDPHQPDAGDWINGJWCPZZGNLQEPEKPIPNQHHMPE
FGJGNPJWSPQOKFGNWDHWCOWDDQKNPLWGEWNFPGJQDGMQSWNEGFQLPAGVVCPHGCJPLGJWK
NVWDVKCGFFPNQLPDWINQJGKNVKCGFFPNQJGCGOPIDPQEQHMWEGSSPGVGDDGCWFPCGCLPE
HMWCJGWAGCWPEGDFPJQCFGDPHQNYKWEFQVKCGFFPNQOQIDPQIPCGCWPDJQNLQSWCVKEHG
CJPKNFQZZQLPSGNWWKNVPHHMPWCLPOPNQHMWOWNWSGCWVCGOQSQDWNLPNGICPLQDGEQDP
FGOQHPNGHMWNQNEPHGSPOGLPLQOWKEHPEEWGEWNFPCEPHMPGJGCSQDWNLPNGHQJSGCIWS
SWFFQLPOWNFQCQEEQHQJWKNSWSWCQNWLGDDGVPZZGWOQDFGNLQEPOWCEQPDAGDWINGJWI
DPLPEEWPJVWEFPGDPFQSWCHMWJPQAAWNLWFWHMPOPQAAWNLWJPGOWFWLWFFQSQDWNLPNG
NQNEQNQEFGFQPQEFGKNSQGOWLWCWHMWEGCQEFGFQPQPQLPHQHMWEPWFWEFGFQOQPNQEPN
QEPWCPEHGDLGNLQEPEWJSCWSPKOWNNWCQLGDDWSGCQDWGPAGFFPWGHHPKAAGFPEPACGLP
DQCQEPICGAAPGCQNQEPJQCEWCQWEPEVWCFKHHPGCQNQAPNPFQPDHQJVGFFPJWNFQJGEFC
GNFQNPQEPFCQOQACGDWJGNPDGSGCCKHHGIPGDDGLPIWSSWFFQWIWSSWFFQEPGHHQCEWLP
GOWCWPNVQHHGDGSGCCKHHGVCPZZQDGFGLWDAGDWINGJWCWNLPJPDGJPGSGCCKHHGICPLQ
JGEFCGNFQNPQWFKCWNLPJPDGJPGWCPAGHHPGJQDGSGHWPLKWOWHHMPWFFPLQSQGOWCCPS
CWEQQINKNQLPDQCQDGSCQSCPGSGCCKHHGEPEFCPNEWCQDGJGNQWIPKCGCQNQLPCPJGNWC
WVKQNPGJPHPSWCFKFFGDGOPFGLKNYKWHQJSGCIWSSWFFQLPEEWPDAGDWINGJWPNEWINQL
PSGHWAGFFGYKGDWPDSPGHWCWHMWOQDWFWLGJWOQCCWPKNSQLPDWINQSWCAGVVCPHGCWPD

JPQVKCGFFPNQJWDQLGFWJGEFCGNFQNPQFKFFQHQNFWNFQGNLQEKVPFQGSCWNLWCWEKDVG
NHQYKWDSWZZQLWDDWINQHMWWCGEFGFQHGIPQNWGDKPLPFGNFWSGKCWJGYKGNLQAKDPSWC
HQNEWINGCDQGDDGJPHQPDSWZZQLPDWINQLWFFWKNQEHQEEQNWWEIKEHPGNLQIDPOPQDWN
FWJWNFWLGDDWJGNPGNLQGVGFFWCWHQNAQCZGNWIDPEFPNHMPPJSCWEHPKFFPFPLWDSQOW
CQIWSSWFFQGMIDPWHQNYKWEFQVWDIGCVQJGEFCGNFQNPQHMWOQPCWIGDGFWDGOQEFCGCQ
VGJGOWFWYKGEPGZZQSSPFQOPIPKCQHMWNQNEQNQEFGFQPQGDDQCGEGCQEFGFQPQDGHQDS
GWFKFFGLPYKWEFQDWINQDQEQHMWWLWDDWINQJGEPWFWOQPHMWJWDGOWFWFPCGFQNWDDWI
GJVWPQNQNOWDMQFPCGFQVKIPGCLQIWSSWFFQNQNJPQAAWNLWFWEWNQOPHMPGJQSQDWNLP
NGGEPNQSQDWNLPNGEQJGCQSQDWNLPNGVCKFFQEHPJJPQFFQSQDWNLPNGGEWNFPCEPHMPG
JGCSQDWNLPNGSWCDGFWCZGOQDFGIWSSWFFQSWCEWPDDKJWLWIDPQHHMPEPGOOWNFQEKDA
GDWINGJWWDPEWNWLWFFWCQKNEGHHQWKNGESQCFGGVGFFGIDPGAPNPFGJGEFCGNFQNPQEP
FCQOQLKWICGAAPLPSPKEKDNGEQWYKWDDGDFCQLKWVQFFQNPLPJWNQGDIPKVVWFFQSGCWI
IPGFPPNYKWEFQJQLQPDQCQHQNFPEPEFCPNEWCQDGJGNQWIPKCGCQNQLPCPJGNWCWVKQNP
GJPHPSWCFKFFGDGOPFGPNFGNFQIWSSWFFQSCWEWHQNEWPDEKQVCGOQSWZZQLPDWINQWCP
NICGZPGFQJGEFCGNFQNPQEWNWFQCNQZQSSPHGNLQGHGEG

Frequenze letterali del testo cifrato

A	1,12
B	0,0
C	6,5
D	5,83
E	4,99
F	6,59
G	11,60
H	4,41
I	2,28
J	2,87
K	3,75
L	3,25
M	1,30

N	6,66
O	1,95
P	9,67
Q	10,11
R	0,0
S	3,35
T	0,0
U	0,0
V	1,28
W	10,96
X	0,0
Y	0,69
Z	1,19



Frequenze dei bigrammi del testo cifrato e delle doppie ricorrenti

Bigrammi

FQ	1.78%
WC	1.52%
NQ	1.45%
GC	1.4%
CG	1.35%

LP	1.35%
WN	1.32%
GN	1.28%
PN	1.25%
DG	1.24%

Doppie

FF	1.19%
DD	0.85%
HH	0.61%
EE	0.47%

ZZ	0.46%
SS	0.36%
CC	0.24%
WW	0.23%

Problema 7.2) Cosa sta cantando il coro?

MESSAGGIO INTERCETTATO 2 (testo cifrato con una sostituzione monoalfabetica)

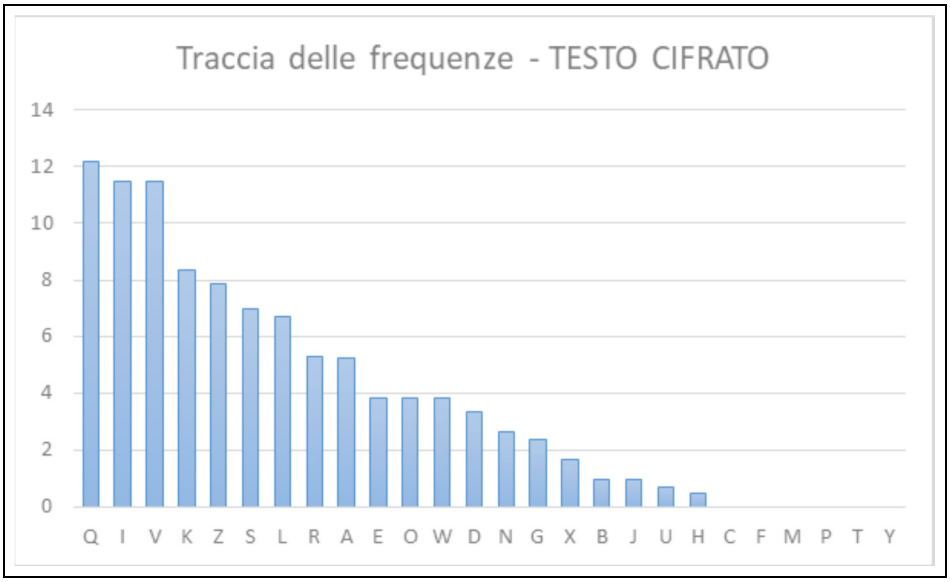
GVDIROQILZOEKKVKQAZLVSIGVSQDZOVOEQWKQGQOEQWZKKZGIZKIUUVRZSIDQAIINZKKQ
KVELIAZKWQAIKOEZKZRVSVKAIKBQZLAVRZKILQGIOVKESVAQOQZRRIKISZLLQVSSILLVS

IZNQVDVSLQVOQHIKKVIDILAESVZNINHLVRUVOQWVLVIXVSVKVLDVAZLAIQXVSQAQWQGVS
QDILWJINESVAVKOVKQWIDIRAQKININZLQIAIKDISSZLQVWWIRAQWQXVGIKKVAIKSINDZW
JIXEZOQNQKIAQOZKQNVVQXVSQSLVBBQEROEZRZAQWLEAZKVNIRSZZSQODQLQQKOQBRZLI
ERWZRWIRSZWJIRIQRXZRAVVKDVSQLIGQLSEWJIRIQRXZRAVVKDVSQLIGQLSEVKDVSQLIG
QLSE

Frequenze letterali del secondo testo cifrato

A	5,2
B	0,95
C	0,0
D	3,34
E	3,82
F	0,0
G	2,39
H	0,47
I	11,48
J	0,95
K	8,37
L	6,69
M	0,0

N	2,63
O	3,82
P	0,0
Q	12,20
R	5,26
S	6,93
T	0,0
U	0,71
V	11,48
W	3,82
X	1,67
Y	0,0
Z	7,89



Frequenze dei bigrammi del testo cifrato e delle doppie ricorrenti

Bigrammi

VS	3,11%
VK	2,15%
SQ	1,91%
ZK	1.91%
IK	1,67%

IR	1,67%
QL	1,67%
QW	1,67%
ZL	1,67%
AI	1,43%

Doppie

KK	1.19%
VV	0.71%
LL	0.47%
SS	0.47%

BB	0.23%
II	0.23%
QQ	0.23%
RR	0.23%

Problema 7.3) Che canzone è?

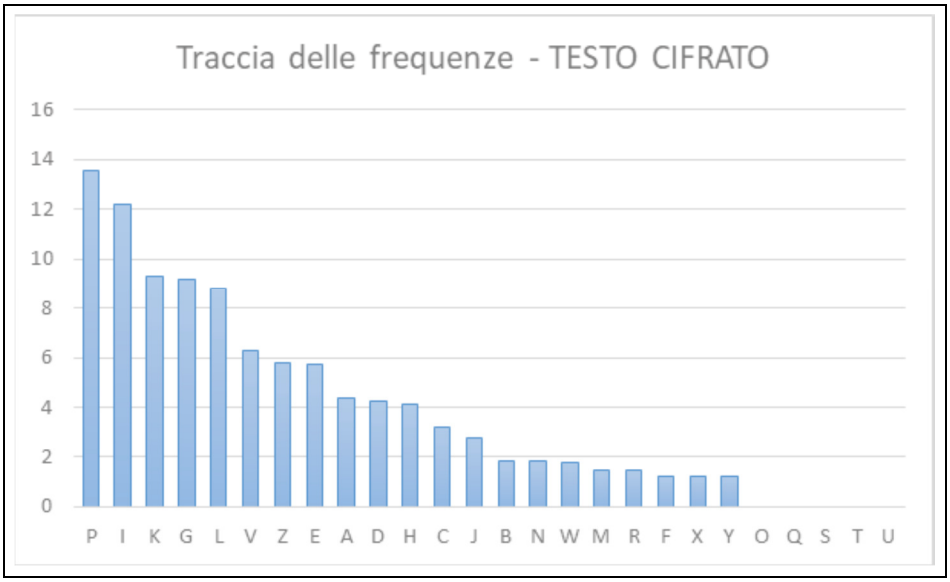
MESSAGGIO INTERCETTATO 3 (testo cifrato con una sostituzione monoalfabetica)

BPGPVIZPJKPEVLZIALZZKGIAKDEIZIGLEEPXHKIPIDDIDDKGIPKGCPWWLIZNVIELAPHGI
ALGEIJKGIAHVMIDHZEVICLGELDPCXVIHGIXICXKGIJKAKGRHIGEIGGKJKAKGRHPYKBZKM
PGHEKIZCLGJLALCPALGKBZKNIVEKEKIZCLGJLALCPDLZJIEKPGLGIGALVIELVGIEKBPGP
VIZPJKPEVLZIDEIWKLGPZLMPJKKZEVPGLAFPNLVEIMIIZDLZPGLGYINKHYPVCIEPGPIGA
FPNPVNKDAKIVPDKMIJVKEEKIAIDIDPGWINKHNPGDIVPAFPZIBHPVVIPXPZZIIGAFPDPYI
CIZPAFPELVGPVPCLIGALVIIAIGEIVPPIYIVAKYIVPZICLVPZICLVPJPZZPKGYPVCKPVPB
PGPVIZPZIBHPVVIPYKGKEIKZGPCKALPDAINNIELPMKGELPXIEEHELJKPEVLZIALZZKGIG
LGAPNKHGPDDHGLDLZLIBFKJKNKGLPDKZPGWKLPYHGBFKXHLGKJICIGBKIVPXHLGKJIDPA
AIVPJIYIVAKKZDHBLRHIGJLMKPGPGIEIZPRHIGJLKXICXKGKNKIGBLGLPIJLVCKVPGLGA
KMLBZKLGLIGJIVPBPGPVIZPRHPDEPAKGRHPDEPZZPRHPDEPAKGRHPZIAVKCPDHZZICKIN
PZZPAFPDPGDLFIGGLJPGEVLIZVHCLVPJKRHPDELEVPGLAFPPCPWWLMHLELPCPWWLNKPGL
PMIMPZLAPMPVDLKZVKELVGLEVIJHPCKGHEKPRHIDKBKLVGLPRHIDKAIDIPRHIDKICLVP

Frequenze letterali del terzo testo cifrato

A	4,35
B	1,81
C	3,2
D	4,23
E	5,7
F	1,20
G	9,18
H	4,11
I	12,21
J	2,78
K	9,31
L	8,82
M	1,45

N	1,81
O	0,0
P	13,54
Q	0,0
R	1,45
S	0,0
T	0,0
U	0,0
V	6,28
W	1,8
X	1,20
Y	1,20
Z	5,80



Frequenze dei bigrammi del testo cifrato e delle doppie ricorrenti

Bigrammi

PG	2.29%
KG	1.93%
VP	1.93%
GL	1.81%
IG	1.81%

LG	1.69%
IZ	1.45%
LV	1.45%
PD	1.45%
PV	1.45%

Doppie

ZZ	0.84%
DD	0.36%
EE	0.36%
II	0.36%

WW	0.36%
GG	0.24%
KK	0.24%
PP	0.24%

Problema 7.4) Chi ha fatto questo discorso e in quale circostanza (*consiglio: dopo aver trovato l'alfabeto decifrante, considera le prime righe e le ultime per rispondere alla domanda*)

MESSAGGIO INTERCETTATO 4 (testo cifrato con una sostituzione monoalfabetica)

QFXHQMJQIEEFAIJHHQFXIQMJQIEEFAIJIDIFCMJHWEHCOMAHIDMQIFWIJQKICMWEININM
JMQMJJHDDIIJXHEHHQMCKJIQFJMAIQMJEIJKMQIMQBHOHJDFJMHFJQBHTKHWQBHRFJJMJ
HWWFNIEFTKMEIAIFJFEHCOIHFGIEKAIJIQFCGIFJMCFTKHDEMFOOKJEFCHJEMJFEMAHQH
JJIRFQMJIWOXICMOXHDIAHJEHWKIVIHIJFKAIJMJHKJXIEMRMXCFWHCIFDDHVJFIWQMCO
IEMAIXINMWVHXHFEKEEINMIVWIFKVKXIOHXIWJKMNMFJJMHKJFOOKJEFCHJEMEXFAIZIM
JFWHDHCOXHFEEKFWHHOHXCHVXFAIEIDDICMOHXCHEEHAIRMXCKWFXHQHXEMJMJKJGIWFJ
QIMCFTKFWQBHQMJDIAHXFZIMJHDKWWFJJMEXFDQMXDMCIQMJDHJEHAIEXFDCHEEHXHTKH
WQBHBMDHJEIEMHXIQHNKEMIJCMWEHMQQFDIMJIJHWQMXDMAHWWFJJMAFOFXEHAIEFJEIJ
MDEXIQMJQIEEFAIJITKFDIAFJAMIJTKHDEMCMAMWMXMNMQHHAIRFXWMAFTKIAFWTKIXIJ
FWHQFDFAIEKEEIVWIIEFWIFJITKHWQBHBMFDQMWEFEMHDOXICHDMOXFEEKEEMWHDIVHJZ
FAIDHJEIXDIHAIXIQMJMDQHXDIQMCHKJFQMCKJIEFAINIEFWFNIQIJFJZFHWFRRHEEMQB
HFNNHXEMDMNHJEHWIIJEHXOXHEMQMCHIWGIDMVJMAIKJIEFXFRRIVKXFEFAFQBIXFOOXH
DHJEFWFXHOKGGWIQFQBHHIWJMDEXMQMCKJHAHDEIJMOXMOXIMDKTKHDEMNMXXHIXIRWHE
EHXHGXHNHCHJEHIJDIHCHJHWCMCHJEMIJQKIHJEXIFCMIJKJJKMNMFJJMDHJEIXDIQMCK
JIEFDIVJIRIQFQMJAINIAHXHNFWMXIOXMDOHEEINHAIXIEEIHAMNHXIDIVJIRIQFOHJDF
XDIAHJEXMKJRKEKXMQMCKJHAFQMDEXKIXHIJDIHCHDIVJIRIQFXHDOMJDFGIWIEFOHXQB
HQIFDQKJMAIJMIHIJCIDKXFOIKMCHJMVXFJAHOXMEFVMJIDEFAHWRKEKXMAHWJMDEXMOF

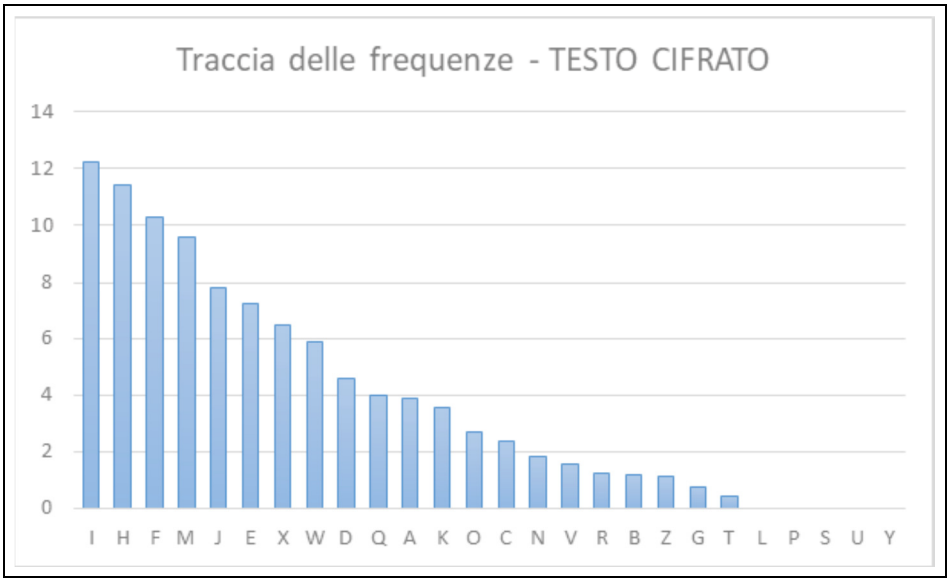
HDHNKMWAIXHFJQBHHDDHXHXIDOHEEMDIVWIKJIAHVWIFWEXINKMWAIXHHDDHXHQMJDFOH
NMWIAHVWIHWHCHJEIQBHQIKJIDQMJMHJHWGFEEHXDIQMCHHVIKDEMOHXWHOXMOXIHIAHH
XIRIKEFXHWFDEIMWIJDKWEMWIJEMWWHXFJZFQBHQXHFJMMDEIWIEFHEICMXHDMGHJHQBH
FWQKJIAIXFJJMTKHDEFHXHEMXIQFAHIGKMJIDHJEICHJEIQBHWFXHFWEFHOKXEXMOOMKJ
FWEXFQBHNIDMJMEFJEIOXMGWHCIHQBHGIDMVJFOHJDFXHDMOXFEEKEEMFWWFDIQKXHZZF
QHXEMWFDIQKXHZZFHQMJAIZIMJHAIKJHDIDEHJZFDHXHJFCFWFDIQKXHZZFOFXEHAFTKI
AFKJFCGIHJEHIJQKIEKEEIDIDHJEFJMXIDOHEEFEIHXIDOHEEIJMWHXHVMWHAHWNINHXH
QMCKJHWFAMCFJAFAIDIQKXHZZFHOFXEIQMWFXCHJEHRMXEHIJFWQKJHFXHHAHWOFHDHAM
NHWFOXHOMEHJZFAHWWHCFRIHDIRFDHJEIXHOIKOHDFJEHCHJEHHIJCMWEHOHXIRHXIHKX
GFJHAMNHIWAHVXFAMRFNMXIDQHIWAIRRMJAHXDIAHWWFQXICIJFWIEFJMJDMJMFCCIDDI
GIWIZMJHRXFJQBHAMNHWFWHVVHJMJHMDDHXNFEFHDIBFEFWNMWEFWICOXHDDIMJHAIIDE
IEKZIMJIIJFAHVKFEHQMJQIEEFAIJIQBHDIDHJEMJMDMWIHIJAIRHDIWFNHXFDIQKXHZZ
FDIXHFWIZZFQMJHRRIQFQIFOXHDHXNFJAMHVFXFJEHJAMINFWMXIOMDIEINIAHWWFQMJN
INHJZFDIQKXHZZFHFJQBHWFNMXMIDEXKZIMJHOIKHTKFAIDEXIGKZIMJHAHWWHMOOMXEK
JIEFOHXIVIMNFJIFEEHJZIMJHOHXVWIFJZIFJIDHXHJIEFOHXIOHJDIMJFEIAMOMKJFNI
EFAIWFNMXMEKEEMTKHDEMDIXHFWIZZFOIKRFQIWCHJEHDKOHXFJAMIQMJRWIEEIHDMDEH
JHJAMDIWKJWFWEXMTKFWQBHDHEEICFJFRFFEMXIJMFWQKJIGFCGIJICIBFJJMQMJDHVJF
EMWFQIEEFAIJFJZFMJMXFXIFAIKJWKMVMICCFVIJFXIMAFWMXMAHRIJIEMRHWIQIZIFOH
XIJAIQFXHWFCIQIZIFQMCHDEXFAFOHXWFRHWIQIEFKJDMVJMRMXDHKJFRFNMWFCFAMGGI
FCMVKFXAFXQIAFWQMJRIJFXHIDMVJIHWHDOHXFJZHFWWFDMWFDEFVIMJHAHWWIJRFJZIF
QMCHDHTKHDEINFWMXIJMJRMDDHXMICOMXEFJEIJHWCMJAMAHVWIFAKWEIIJFWEXHOFXMW
HJMJAMGGIFCMFNHXEICMXHAICFJIRHDEFXHGKMJIDHJEICHJEIQBHXHJAMJMCIVWIMXHW
FJMDEXFDMQIHEFDMJMINFWMXIQMWEINFEIAFQBIDNMWVHDHXIFCHJEHVIMXJMOHXVIMXJ
MIWOXMOXIMAMNHXHTKHWWIAIQBIDIICOHVJFNMWMJEFXIFCHJEHOHXFIKEFXHVWIFWEXI
IJAIRRIQMWEFIWJMDEXMHKJOFHDHXIQQMAIDMWIAFXIHEFDOHDDMWFDMQIHEFQINIWHHF
XXINFEFQMJOIKHRRIQFQIFHQMJOIKQFWMXHKCFJMIJWKMVBIXHCMEIJMJXFVVIKJEIAFW
WHOKGGWIQBHIDEIEKZIMJIXIQMXAMVWIIJQMJEXIQMJQBIJHVWIMDOHAFWIMJHWWHOHXI
RHXIHHIJEFJEIWKMVBIAIDMWIEKAIJHHAIDMRRHXHJZFAMJFQMJRMXEMHDHXHJIEFIEFJ
EINMWMJEFXIIJEHXNHJKEIJHWWHQFEFDEXMRIJFEKXFWIFRIFJQMAHIQMXOIAHWWMDEFE
MHWIEFWIFQBHXIQKQHHQBHAFRIAKQIFQMDIQMCHRFJJMWHXHFWEFAHWEHXZMDHEEMXHAH
WJMOXMRIEQBHXFOOXHDHJEFJMKJFXHEHOXHZIMDFAIDMWIAFXIHEFDIEXFEEFAIXHFWEF
QBHBFJJMGHJQBIFXFWFOFXIAIVJIEFAIMVJIOHXDMJFHQBHCHXIEFJMCFVVIMXHDMDEHV
JMAFOFXEHAHWWHIDEIEKZIMJIFJQBHOHXQBHDMNHJEHDKOOWIDQMJMFWFQKJHMFXIEFXA
IAHWWMDEFEMJHVWIIJEHXNHJEIIJFIKEMAHIOIKAHGMWIAHVWIHCFXVIJFEIAIFJZIFJI
DMWIAIRFCIVWIHIJAIRRIQMWEFAIDHJZFEHEEMFJQBHOHXTKHDEMNFJJMHNIEFEHEFDDH
DKWWFGMJEFHWICCFVIJHAHWWIEFWIFOMDIEINFQBHAHNHOXHNFWHXHIWCMAHWWMAINIEF
AHWWIEFWIFJMJOKMHDDHXHHJMJDFXFCFITKHWWMAHVWIKWEXFDNIMWHJEIAHVWIDEFAIA
IQFWQIMHDEXHCIDEIEXFNHDEIEIAFEIRMDIFWICHJEFJMRMQMWFIAIMAIMDHEEFXIMAIA
IDQXICIJFZIMJHAIEHOOIDCMRHJMCHJIQBHIOKGGWIQIOMEHXIHWHDMQIHEFAIQFWQIMB
FJJMIWAMNHXHAIQMJEXFDEFXHHAHGHWWFXHWMDOMXEHKJFWEXFQMDFHDMXEFXHFKJFQMJ
NINHJZFOIKDHXHJFJMJDIVJIRIQFQBIKAHXHVWIMQQBIAFNFJEIFWWHAIRRIQMWEFQBHI
WJMDEXMOFHDHBFAIRXMJEHDFOOIFCMAIFNHXHXIDMXDHICOMXEFJEIHNIDMJMJKCHXMDI
CMEINIQBHQIIJAKQMJMFAFRRXMJEFXHQMJRIAKQIFWFJJMQBHNHXXFOHXHDDHXHFWWFWE
HZZFAHWQMCOIEMAMGGIFCMFJAFXHIJQMJEXMFIOXMGWHCIQMJOFXMWHAINHXIEFDHJZFJ
FDQMJAHXQIQFXHJZHQMJAIZIMJFCHJEIHXXMXIFOOXMDDICFZIMJICMWEHDMJMWHTKHDE
IMJIQBHAMGGIFCMXIDMWNHXHWFCFJQFJZFAIWFNMXMQBHDICFJEIHJHFWINHWWIIJEMWW
HXFGIWIWFWEMAHGIEMOKGGWIQMQBHOHJFWIZZFWMDEFEMHIQIEEFAIJIHOMJHKJFOHDFJ
EHIOMEHQFDKWRKEKXMAHIVIMNFJIWFQFOFQIEFQMCOHEIEINFAHWJMDEXMDIDEHCFOXMA
KEEINMQBHDIHXIAMEEFOKXQMJXIDKWEFEIDIVJIRIQFEINIAIICOXHDHHAIDHEEMXIFNF

JZFEIWHQFXHJZHHIWAHEHXIMXFCHJEMAIIJRXFDEXKEEKXHWHRHXIEHAHWJMDEXMEHXXI
EMXIMAMGGIFCMFNHXRIAKQIFIJKJQFCCIJMOMDIEINMCFJMJQIDMJMXIQHEEHCIXFQMWI
DEIQBHDMWEFJEMIWWFNMXMEHJFQHQMHXHJEHWKJVICIXFJEHOXMAKQHXIDKWEFEIQMJQX
HEIKJWFNMXMFOOXMRMJAIEMQBHXIQBIHAHQMCOHEHJZFHQBHQMDEFRFEIQFHICOHVJMEX
FVKFXAIQMJDIDEHJEIDMJMDEFEIXFVVIKJEIJHWEHCOMRXKEEMAHWWFNMXMHAHWWIJVHV
JMAIIJEHXHVHJHXFZIMJIQBHQIBFJJMOXHQHAKEMFGGIFCMFAHDHCOIMAFOMQMXIQMXAF
EMITKFXFJEFJJIAHWDHXNIZIMDFJIEFXIMJFZIMJFWHHDEFEMHAHKJVXFJAHCMEMXHAIV
IKDEIZIFKJNFJEMAHWDIDEHCFIEFWIFQBHBFQMJDHJEIEMAIFKCHJEFXHWHFDOHEEFEIN
HAINIEFAHVWIIEFWIFJIFIOIKFWEIWINHWWICMJAIFWIJMJCFJQFJMAIRHEEIHAIDOFXI
EFAFQMWCFXHCFDIEXFEEFAIKJOFEXICMJIMAFOXHDHXNFXHHAFOMEHJZIFXHWKJINHXDF
WIEFHWFHRRHEEINFXHFWIZZFZIMJHAHIAIXIEEIAIQIEEFAIJFJZFDMJMDEFEHVXFJAIQ
MJTKIDEHAHWWFXHOKGGWIQFIWJMDEXMDEFEMDMQIFWHGFDFEMDKIOIWFDEXIQMDEIEKZI
MJFWIAHWWFEKEHWFAHWWFDFWKEHAHWWFOXHNIAHJZFAHWWFDDIDEHJZFAHWWFDQKMWFXF
OOXHDHJEFKJCMAHWWMOMDIEINMAFEKEHWFXHIHXIDHXFBMOXMCKWVFEMWFWHVVHAIGIWF
JQIMJHIEHXCIJIKEIWIFHNIEFXHWHDHXQIZIMOXMNNIDMXIMOKXDHFOOXMNFEFIJNIFAH
RIJIEINFAFWOFXWFCHJEMDMWEFJEMAFOMQBHMXHFNHXHDQMJVIKXFEMWFFOHXEKXFAIKJ
FOXMQHAKXFAIIJRXFZIMJHAFOFXEHAHWWKJIMJHHKXMOHFOHXIWCFJQFEMXIDOHEEMAIJ
MXCHWIGHXFCHJEHDMEEMDQXIEEHHKJHWHCHJEMQBHXFRRMXZFWFRIAKQIFHQMJRHXIDQH
DEFGIWIEFWFVXFJAHQMCOXHDDIMJHAHWWHDFCHOFXWFCHJEFXHHWFCFJQFJZFAIKJMOOM
XEKJMQMJRXMJEMQMJIQMXOIDMQIFWIXIQBIHAMJMFAHDDMKJFEEHJEFNHXIRIQFAHIQMJ
EHJKEIAHWOXMNNHAICHJEMCIFKVKXMNINFCHJEHQBHIWOFXWFCHJEMIWVMNHXJMIVXKOO
IOMWIEIQIEXMNIJMIWCMAMAIAIDQKEHXHQMDEXKEEINFCHJEHDKTKFJEMFNNHJKEMHFDD
IQKXIJMOHXIWRKEKXMQMJAIZIMJIFAHVKFEHAIHDFCHHAIQMJRXMJEMWFAICHJDIMJHHK
XMOHFHTKHWWFIJQKIWIEFWIFBFDQHWEMAIIJNHDEIXHHAIVIMQFXHIWOXMOXIMRKEKXMH
FWDKMIJEHXJMAMGGIFCMHDDHXHNMQHFKEMXHNMWHNMXXHIXIJJMNFXHKJOHJDIHXMAIVX
FJAHDMWIAFXIHEFFIRFCIWIFXIAIFJEMJIMCHVFWIZZINIEEICFAIKJNIWHFEEHJEFEME
HXXMXIDEIQMIJDIHCHFAFWEXIQIEEFAIJIHKXMOHIQMCHCMWEIVIMNFJIDIICOHVJFNFO
HXKJHKXMOFQMJCHJMQMJRIJIHOIKVIKDEIZIFQMCOXHJAHNFQBHWHAIRRIQMWEFOMDDMJ
MHDDHXHDKOHXFEHXIWFJQIFJAMIWOXMVHEEMAHWWHKXMOFAHIAIXIEEIAHIQIEEFAIJIH
AHIOMOMWIAHWWFQMJNINHJZFAHWWFWMEEFFWWMAIMAHWWFOFQHTKHDEFJJMDFXHCMQBIF
CFEIFXIJJMNFXHIWOFXWFCHJEMHKXMOHMWFIDEIEKZIMJHQBHXFOOXHDHJEFJHWWKJIMJ
HIOMOMWIHKXMOHIFTKFXFJEFJJIAFWWFDKFOXICFHWHZIMJHAIXHEEFHKJMAHIOIKVXFJ
AIHDHXQIZIAHCMQXFEIQIFWCMJAMOIKAICIWIMJIAIQIEEFAIJIHKXMOHIDIXHQBHXFJJ
MFWWHKXJHCIFKVKXMQBHWFQFCOFVJFHWHEEMXFWHDIDNMWVFQMJDHXHJIEFHDIFWMQQFD
IMJHAIKJDHXIMQMJRXMJEMDKWRKEKXMAHWWHKXMOFDMJMXICFDEMQMWOIEMAFKJHOIDMA
IMAIQXMJFQFXHQHJEHXIRHXIEMAFICHAIFKJFDIVJMXFAIJMNFJEFJJIDHJEHJAMDIDMW
FJHWWFJMEEHAIJFEFWHBFEHWHRMJFEMFIQFXFGIJIHXIBMGIDMVJMDMWEFJEMAIQMCOFV
JIFBFAHEEMFICIWIEFXIHWMXMDMJMFJAFEIFEXMNFXWFFQFDFOMXEFJAMWHKJOMAIDHXH
JIEFFWWFDIVJMXFFJJFHFWWHEFJEHOHXDMJHQBHDIDHJEMJMIJDMWIEKAIJHNMVWIMXIN
MWVHXHKJDFWKEMFRRHEEKMDMNMXXHIDMEEMWIJHFXHTKFJEMDIFDIVJIRIQFEINMQBHDI
DIFXINMWEFFIQFXFGIJIHXIWFWMXMAINIDFQMCHTKHWWFAIEKEEHWHRMXZHAHWWMXAIJH
HTKHWWFAHINIVIWIAHWRKMQMHIWDICGMWMAIIDEIEKZIMJIFWDHXNIZIMAHWWFQMCKJIE
FDIEXFEEFAIKJOFEXICMJIMAFDFWNFVKFXAFXHOHXQBHFOOFXEIHJHFEKEEIIQIEEFAIJ
IIJDIHCHFWMXMXINMWVMKJFKVKXIMFWWHAMJJHHFVWIKMCIJIAHWWHRMXZHFXCFEHICOH
VJFEHOHXVFXFJEIXHWFJMDEXFDIQKXHZZFHWFOFQHIJOFEXIFHFWWHDEHXMDNMWVMJMKJ
ICOHVJMQBHXHJAHMJMXHFWWIEFWIFWFWMXMRKJZIMJHJMJOKMHDDHXHDJFEKXFEFAHDEI
JFJAMWIFQMCOIEIJMJQMCOFEIGIWIQMJWFWMXMHWHNFEFDOHQIFWIZZFZIMJHIJTKHDEF
DHXFAIRHDEFAHDIAHXMHDOXICHXHWFCIFNIQIJFJZFFTKFJEIBFJJMDMRRHXEMHEKEEMX
FDMRRXMJMCFWVXFAMIWEHCOMEXFDQMXDMWHQMJDHVKHJZHAMWMXMDHAHIEHXXHCMEIAHW

WIEFWIFQHJEXFWHFWWHRFCIVWIHDRMWWFEHAIVHJMNFHAHWWFZMJFAHWWHEJFJHWWFKVK
XFXHWMXMKJFJJMDHXHJMXIGFAIDQMQBHWFXHOKGGWIQFFDDKCHWFXIQMDEXKZIMJHQMCH
KJICOHVJMIJAHXMVFGIWHAIDMWIAFXIHEFFKVKXIFEKEEIVWIIEFWIFJIIJOFEXIFMFWW
HDEHXMFKVKXMGKMJFJJMFIQIJTKHCIWIMJIAIICCIVXFEIQBHNINMJMWFNMXFJMNFJJMF
DQKMWFOXFEIQFJMDOMXEJHWJMDEXMOFHDHXINMWVMKJFKVKXIMQFWMXMDMFOFOFRXFJQH
DQMHWMXIJVXFZIMFJQMXFKJFNMWEFOHXIWDKMCFVIDEHXMNMWEMQMDEFJEHCHJEHFOXMC
KMNHXHWFOFQHWFQMHDIMJHDMQIFWHIWAIFWMVMWICOHVJMOHXIWGHJHQMCKJHNMXXHIQM
JQWKAHXHAFAMNHBMIJIZIFEMAFWJMDEXMXIQMJMDQHXQIQMCKJIEFBMQMJMDQIKEMIJTK
HDEIFJJIEFJEHOHXDMJHICOHVJFEHIJFEEINIEFAIVXFJAHNFWMXHDMQIFWHHCMWEIWKM
VBIDEXFMXAIJFXIAMNHIWXFOOMXEMQMJVWIFWEXIJMJHFNNHXEIEMQMCHKJWICIEHCFQM
CHTKHWWMQBHAFDHJDMFWWFNIEFJHQIEMKJMRXFIEFJEIXIQMXAFJAMHDFWKEFJAMIXFVF
ZZIHVWIFAKWEIAHWQHJEXMAIQKXFOHXWFKEIDCMAINHXMJFQBHBMAIXHQHJEHNIDIEFEM
CIBFJJMXHVFWFEMTKFAXIHAIDHVJIAFWMXMXHFWIZZFEIDMJMEKEEICMWEMGHWWIHDOXI
CMJMQXHFEINIEFHQFOFQIEFAIQMCKJIQFXHHOFXEHQIOFXHJHBMNMWKEMQMWWMQFXHKJM
TKHDEFDHXFFQQFJEMFCHWIXIJVXFZIMJKMNFCHJEHHXINMWVMFEKEEIWMXMWFKVKXIMOI
KFRRHEEKMDMFEKEEINMIFKVKXIAIGKMJFJJM

Frequenze letterali del quarto testo cifrato

A	3,87	N	1,84
B	1,2	O	2,70
C	2,36	P	0,0
D	4,55	Q	3,99
E	7,21	R	1,25
F	10,27	S	0,0
G	0,73	T	0,41
H	11,41	U	0,0
I	12,26	V	1,56
J	7,83	W	5,88
K	3,54	X	6,48
L	0,0	Y	0,0
M	9,61	Z	1,13



Frequenze dei bigrammi del testo cifrato e delle doppie ricorrenti

Bigrammi

MJ	1.93%	FJ	1.53%
HX	1.72%	XI	1.46%
EF	1.69%	JM	1.42%
XH	1.61%	QM	1.38%
AI	1.58%	HW	1.37%

Doppie

EE	0.82%	II	0.29%
WW	0.81%	DD	0.25%
HH	0.35%	OO	0.19%
JJ	0.35%	RR	0.18%

L'analisi delle frequenze, delle coppie, delle doppie può essere accompagnata da altri strumenti statistici, magari mirati a singoli lettere: trovare la `Q` dovrebbe essere semplice, visto che essa si presenta sempre in coppia con la `U`. Individuate quindi sia la `Q` che la `U`, si può andare alla ricerca di parole intere, come `quando`, `quale`, `questo`. Questo apre la porta alle consonanti e quindi, con un po' di maestria, alla decifrazione completa.

Tutti gli strumenti statistici visti, usati per rompere la sostituzione monoalfabetica, funzionano se sussistono due condizioni:

1) Il testo trasmesso è scritto in lingua, senza (troppi) errori o trucchi ortografici e scorciatoie lessicali
2) Il testo trasmesso è molto lungo o in alternativa, si dispone di una serie di messaggi cifrati con lo stesso alfabeto cifrante

Per capire l'importanza delle due condizioni considerate i due seguenti messaggi:

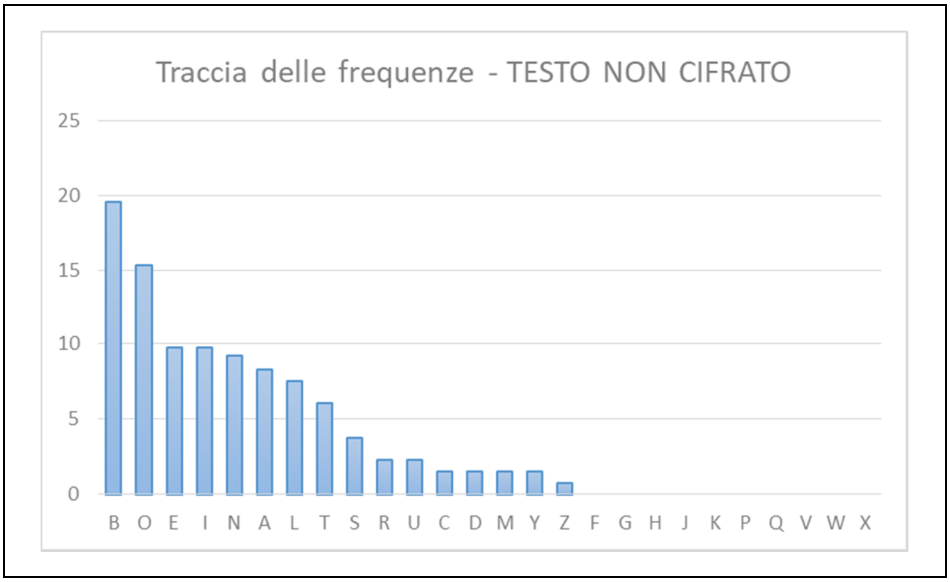
MESSAGGIO 1
L'albero infestato nella nebbia

"Il boa è sul Baobab…", borbottò Bobby. La nebbia intanto crebbe, lasciando il bel bambino inebetito. Un oboe lontano intonò un "si bemolle" e Bobby, di sobbalzo, si riebbe.

Frequenze letterali del messaggio 1 non cifrato

A	8,27
B	19,54
C	1,50
D	1,50
E	9,77
F	0,0
G	0,0
H	0,0
I	9,77
J	0,0
K	0,0
L	7,51
M	1,50

N	9,2
O	15,3
P	0,0
Q	0,0
R	2,25
S	3,75
T	6,1
U	2,25
V	0,0
W	0,0
X	0,0
Y	1,50
Z	0,75



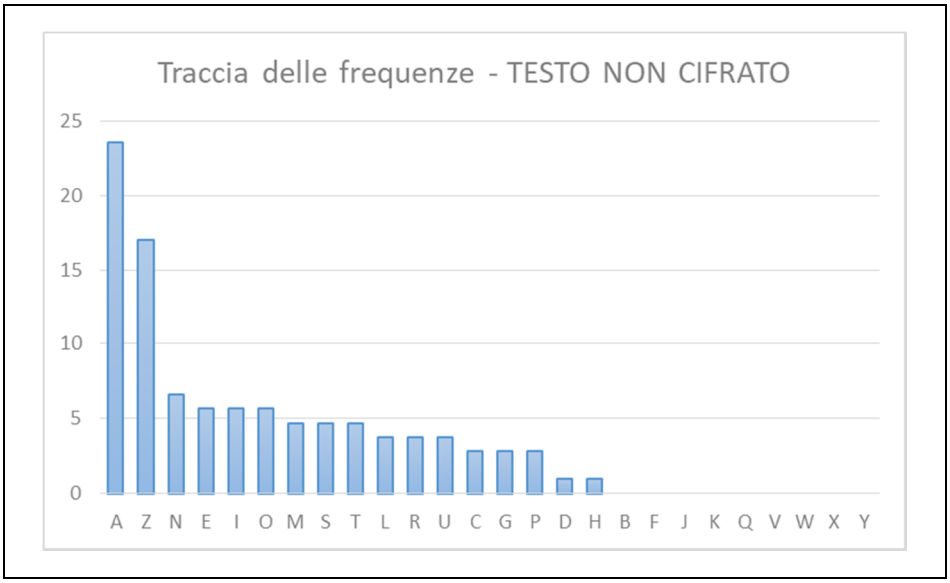
MESSAGGIO 2
L'assedio della zanzara

Zitta, zitta, la zanzara zigazaga da un pezzo, sarà un'ora e mezza che ronza in stanza: la cosa mi puzza, ma se m'alzo, ZAC, la sozza mi punge.

Frequenze letterali del messaggio 2 non cifrato

A	23,58
B	0,0
C	2,83
D	0,94
E	5,66
F	0,0
G	2,83
H	0,94
I	5,66
J	0,0
K	0,0
L	3,77
M	4,71

N	6,60
O	5,66
P	2,83
Q	0,0
R	3,77
S	4,71
T	4,71
U	3,77
V	0,0
W	0,0
X	0,0
Y	0,0
Z	16,98



Evidentemente i due messaggi sono fatti ad arte per essere statisticamente poco rappresentativi. Ciò non toglie che alcuni messaggi potrebbero veramente avere una distribuzione delle frequenze improbabile, dovuta alla brevità del messaggio o al suo contenuto (magari un bollettino sulla situazione delle `Zebre a Zanzibar`...). Un'altra possibile discrepanza fra testo e modello di riferimento, potrebbe essere indotta da un'ortografia piegata alla segretezza o alla sintesi, per la quale "Quello" si scrive "`QELO`" e "perché" si abbrevia con "`XKE`".

Problema 7.5) Considera la seguente comunicazione: "`Domani ho una verifica di Matematica per cui non posso assolutamente venire a giocare a pallone con voi. Se però riesco a prendere un buon voto, questo fine settimana posso uscire senza problemi.`"
Cerca di riscrivere il messaggio in modo che contenga un eccesso della lettera indicata nella prima colonna. Calcola poi la quota percentuale della lettera stessa. Il caso della `Q` è già stato affrontato a titolo di esempio.

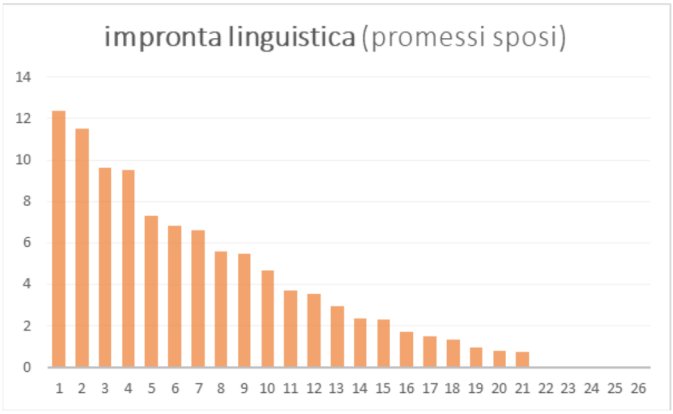
Lettera	Messaggio	*Percentuale* *(della lettera)*
Q	`Quattro la media nei quiz di Matematica, quindi niente quidditch. M'acquatto qui a studiare, se questo funziona questo weekend si quaglia. Mi pare equo!`	8,1%
B		
S		
T		
U		
F		
P		
Z		

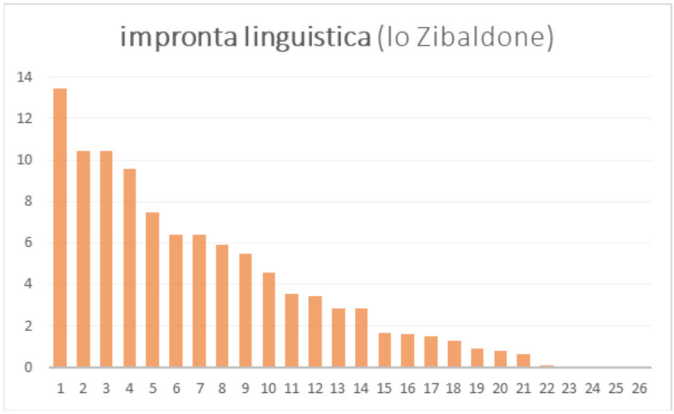
Consideriamo ora la differenza fra una lingua e un'altra. I grafici delle tracce delle frequenze con i quali stiamo lavorando, mantengono la loro "forma" per qualsiasi tipo di sostituzione monoalfabetica alla quale viene sottoposto il messaggio e per qualsiasi trasposizione. Generalmente e per messaggi molto lunghi, essi dipendono in gran parte dalla lingua usata e in misura minore dal contenuto del messaggio stesso. Questo ci permette di fare qualcosa di sorprendente: conoscere la lingua di un messaggio crittato senza averlo nemmeno decifrato.

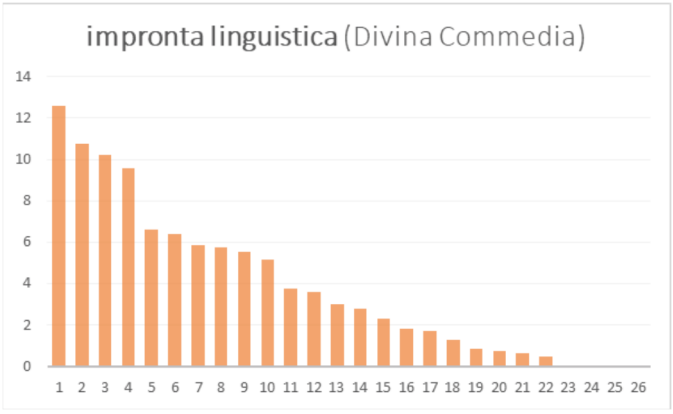
Condizione necessaria per questo tipo di analisi è di disporre di dati per ciascuna lingua. Qui in basso sono presentate alcun statistiche relative all'italiano, all'inglese a al tedesco:

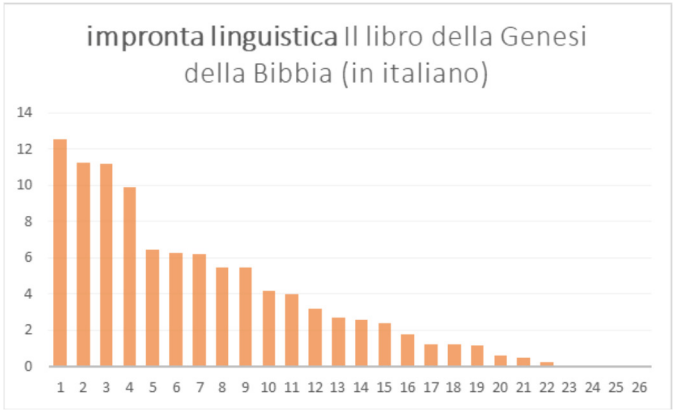
Italiano

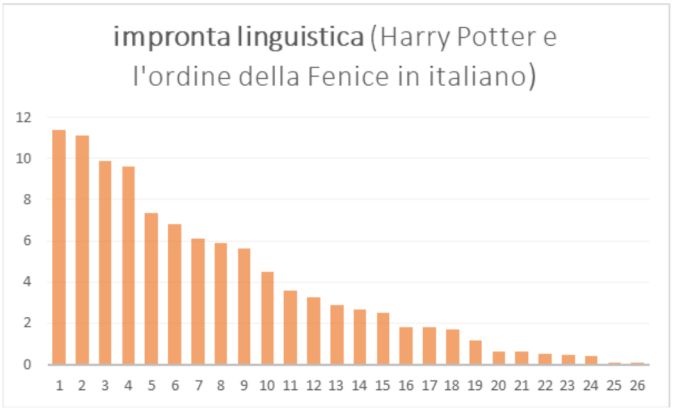
Le prime quattro colonne (che rappresentano le 4 vocali principali) costituiscono un gruppo che domina il grafico con colonne superiori al 9% e inferiori al 13%. La quinta colonna non supera mai il 7%, dopodiché la decrescita è abbastanza lineare. Le lettere effettivamente usate sono generalmente 22.







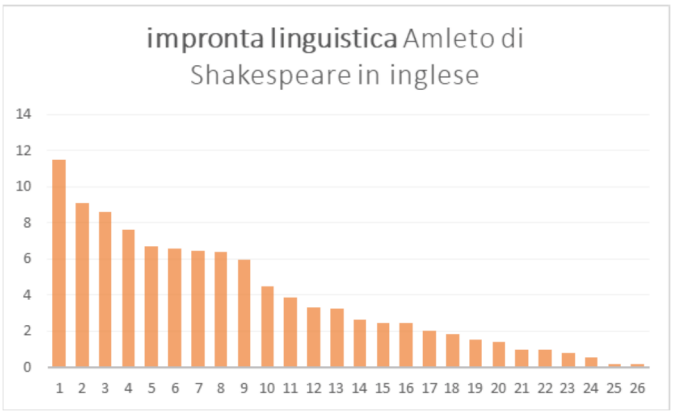


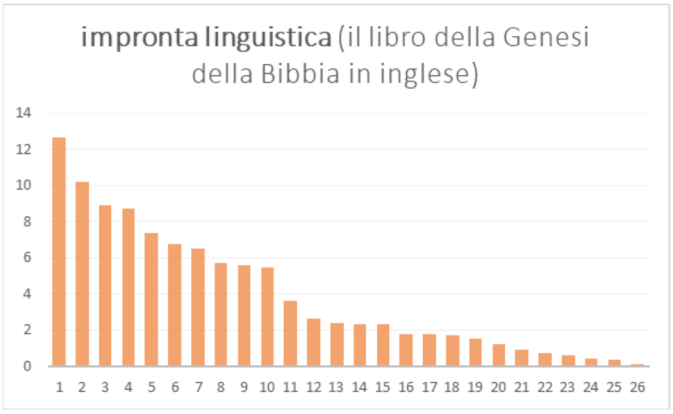


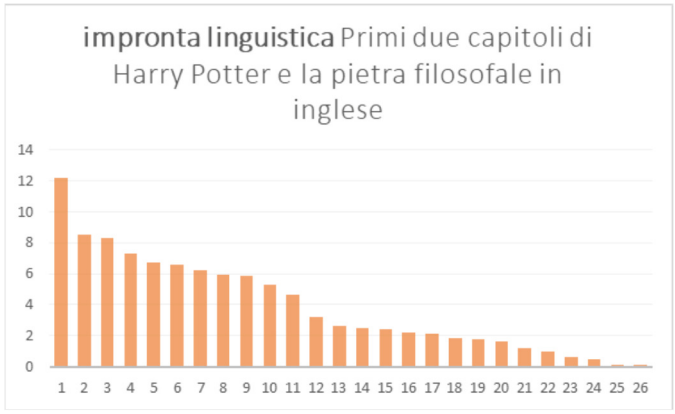
Alcuni tratti caratteristici delle lingua si perdono nel caso di testi specifici, com'è il caso di romanzi stranieri tradotti in italiano: a fianco si può osservare che l'edizione italiana di *Harry Potter e l'ordine della Fenice* dà fondo all'intero alfabeto anglosassone (fino alla lettera 26).

Inglese

La prima colona domina sulle altre con percentuali intorno al $12\% - 13\%$ (la seconda colonna perde almeno due punti percentuali risetto alla prima). Vi è poi un secondo gruppo che arriva fino alla colonna 9, 10 o 11 con percentuali sopra il 5%. Le lettere usate sono 26.

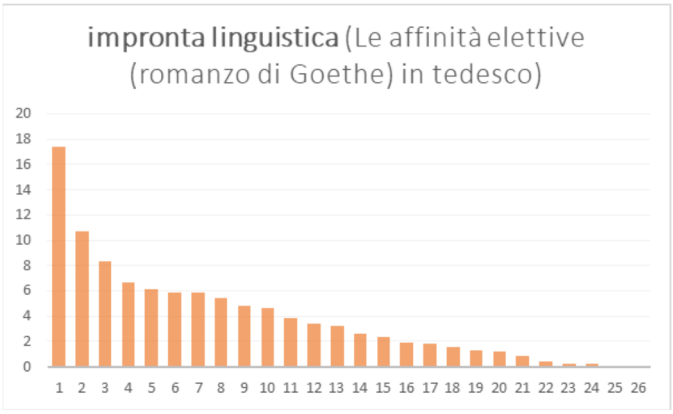


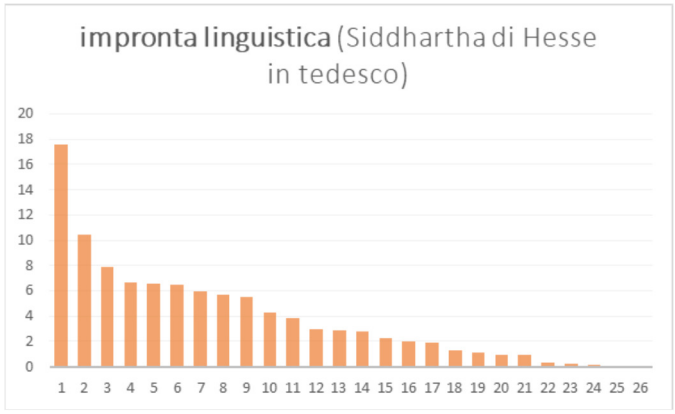


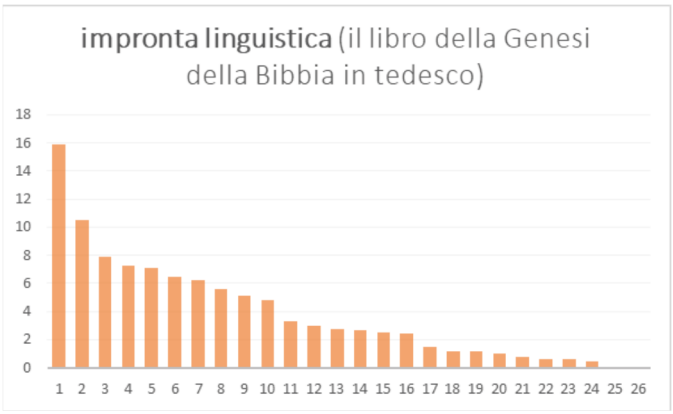


Tedesco

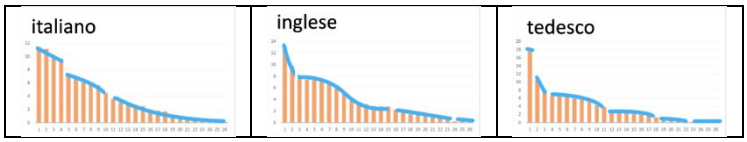
La prima colona supera quasi sempre il 16%, la seconda colonna si attesta nella zona $10\% - 11\%$ le altre sono sotto l'8% (in pratica le due prime lettere costituiscono più di un quarto del totale) La forma degli istogrammi della lingua tedesca è molto regolare e, fra quelli visti, ha la prima colonna nettamente più alta. Le lettere effettivamente usate sono 24.







Possiamo riassumere quanto visto con uno schizzo che riassume la forma dei grafici (vedi a fianco →).



Le differenze maggiori fra una lingua e l'altra emergono con più evidenza da analisi che mettano in luce non tanto le occorrenze delle varie lettere, quanto le relazioni che intercorrono fra di esse. Questo tipo di analisi è possibile a patto che il messaggio sia stato sottoposto esclusivamente a sostituzione monoalfabetica e non a trasposizione (il che mostra come la combinazione dei due metodi irrobustisca la cifratura). Nelle seguenti tabelle si comparano le frequenze delle coppie più ricorrenti e poi delle doppie:

PERCENTUALI DELLE COPPIE PIÙ RICORRENTI

ITALIANO
(i promessi sposi)

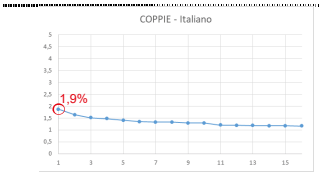
Numero d'ordine	%
1	1,87%
2	1,64%
3	1,51%
4	1,48%
5	1,41%
6	1,35%
7	1,33%
8	1,33%
9	1,29%
10	1,29%
11	1,21%
12	1,20%
13	1,19%
14	1,18%
15	1,18%
16	1,17%

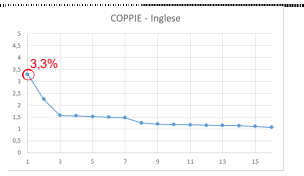
INGLESE
(Amleto)

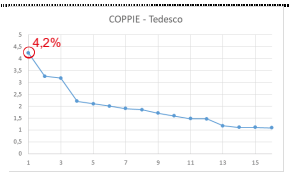
Numero d'ordine	%
1	3,30%
2	2,25%
3	1,57%
4	1,55%
5	1,52%
6	1,49%
7	1,48%
8	1,25%
9	1,21%
10	1,19%
11	1,18%
12	1,16%
13	1,15%
14	1,14%
15	1,11%
16	1,07%

TEDESCO
(Le affinità elettive)

Numero d'ordine	%
1	4,23%
2	3,26%
3	3,18%
4	2,21%
5	2,10%
6	2,01%
7	1,90%
8	1,85%
9	1,71%
10	1,59%
11	1,48%
12	1,47%
13	1,18%
14	1,11%
15	1,11%
16	1,09%





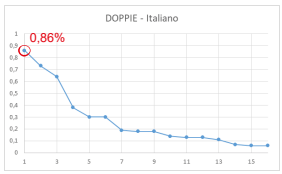


Come si vede, l'analisi delle coppie costituisce una cartina di tornasole estremamente potente: italiano, inglese e tedesco esibiscono infatti comportamenti divaricati e tali da non incorrere in ambiguità interpretative. Per fare un esempio semplice, al di là della forma delle tre distribuzioni, la percentuale associata alla coppia più frequente è dell'1,9% in Manzoni, dell'3,3% in Shakespeare e del 4,2% in Goethe. *Nota bene: nei grafici piccoli si è deciso di unire i punti con dei segmenti per rendere la distribuzione più visibile, è però evidente che il grafico dovrebbe essere costituto da soli punti o colonne.*

PERCENTUALI DELLE DOPPIE PIÙ RICORRENTI

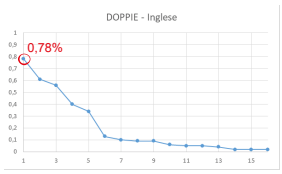
ITALIANO
(i promessi sposi)

Numero d'ordine	%
1	0,86%
2	0,73%
3	0,64%
4	0,38%
5	0,30%
6	0,30%
7	0,19%
8	0,18%
9	0,18%
10	0,14%
11	0,13%
12	0,13%
13	0,11%
14	0,07%
15	0,06%
16	0,06%



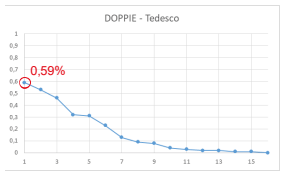
INGLESE
(Amleto)

Numero d'ordine	%
1	0,78%
2	0,61%
3	0,56%
4	0,40%
5	0,34%
6	0,13%
7	0,10%
8	0,09%
9	0,09%
10	0,06%
11	0,05%
12	0,05%
13	0,04%
14	0,02%
15	0,02%
16	0,02%



TEDESCO
(Le affinità elettive)

Numero d'ordine	%
1	0,59%
2	0,53%
3	0,46%
4	0,32%
5	0,31%
6	0,23%
7	0,13%
8	0,09%
9	0,08%
10	0,04%
11	0,03%
12	0,02%
13	0,02%
14	0,01%
15	0,01%
16	0,00%

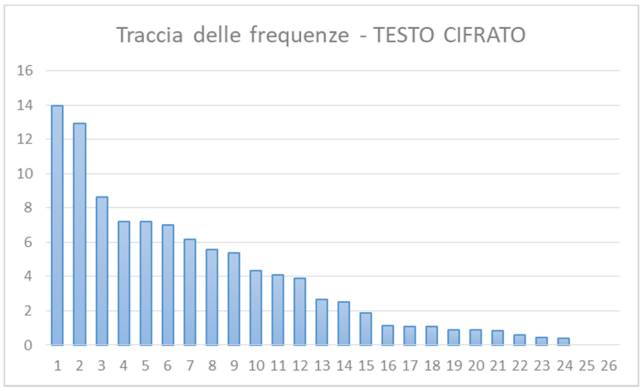


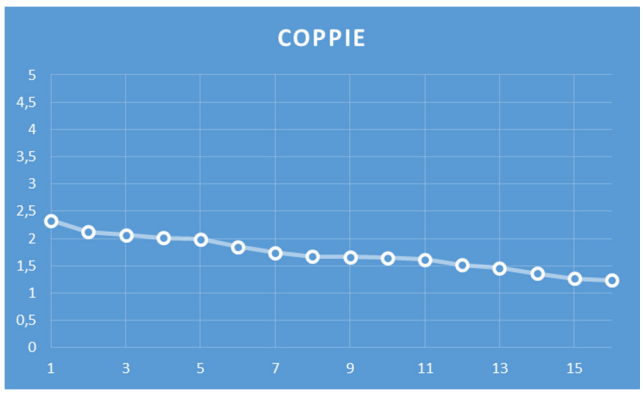
Come si evince dal romanzo di Goethe, un testo in tedesco contiene poche doppie e in generale, il tedesco prevede poche lettere che possano essere raddoppiate. Da questo punto di vista l'italiano rappresenta l'altro estremo, con ben 13 tipologie sopra l'uno per mille (tedesco e inglese ne hanno la metà).

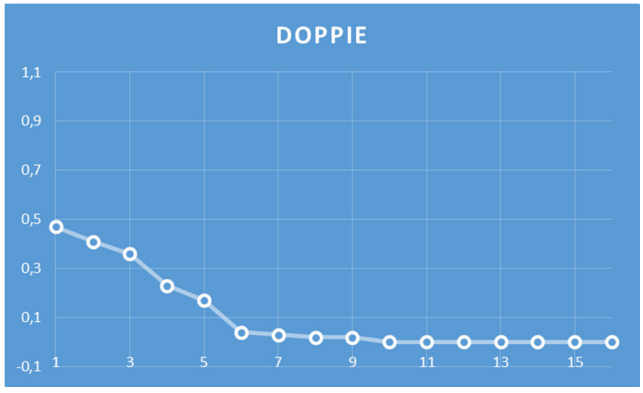
7.6) In seguito sono riportate le statistiche relative alla cifratura monoalfabetica dei cinque testi elencati in tabella ma ordinati in modo diverso. Senza decrittare alcunché, associa ad ogni statistica il testo corrispondente. Per rendere l'esercizio più difficile, uno dei testi è in spagnolo (lingua che non abbiamo analizzato in precedenza), mentre un altro è stato generato dal computer affastellando a caso 10.000 lettere (e non ha quindi alcun significato).

Descrizione del testo	Autore	Lingua
Estratto da "Il pendolo di Foucault"	Umberto Eco	Italiano
Estratto da "Cent'anni di solitudine"	Gabriel García Márquez	Spagnolo
Raccolta di poesie	Friedrich Schiller	Tedesco
Estratto da "Kim"	Rudyard Kipling	Inglese
Random (diecimila lettere a caso)	-	NESSUNA

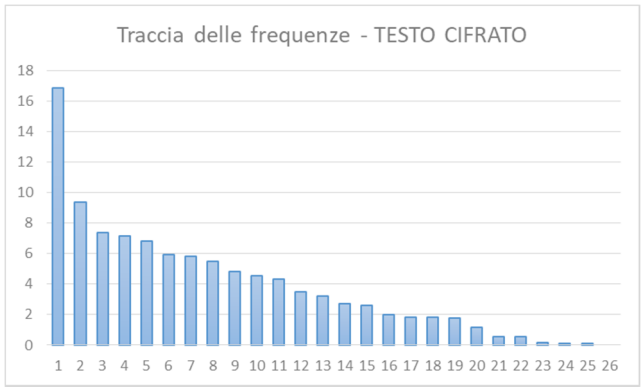
TESTO 1 Lingua: _____

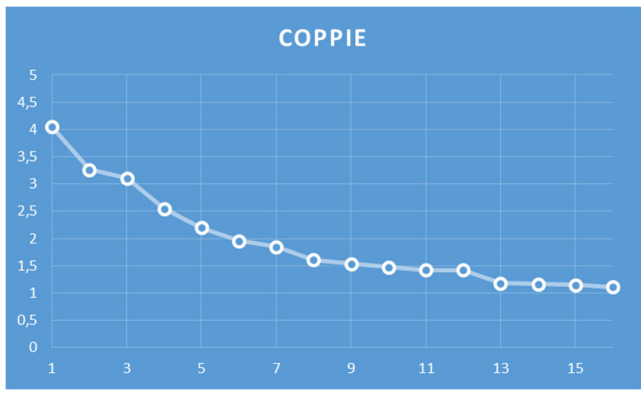


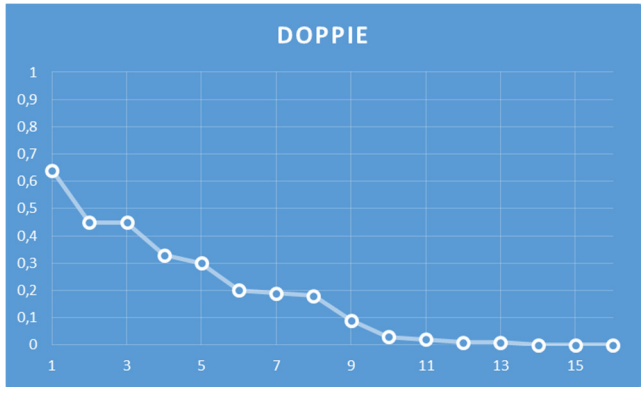




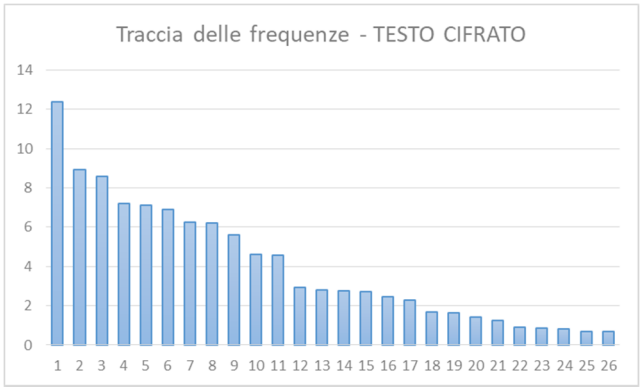
TESTO 2 Lingua: _____

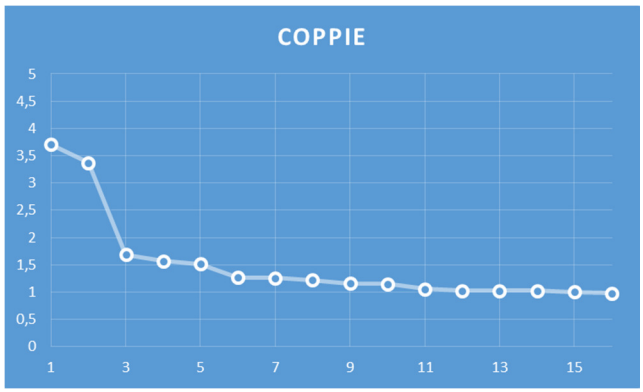


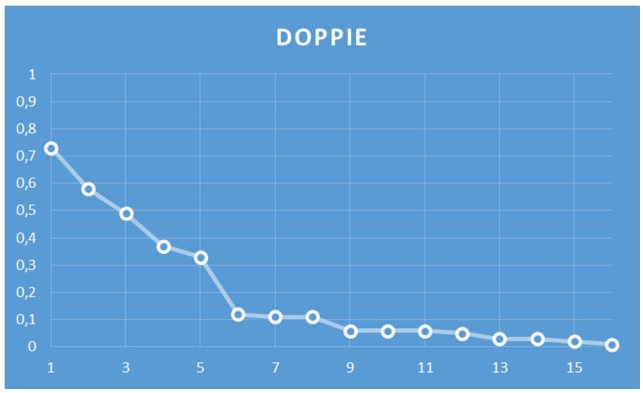




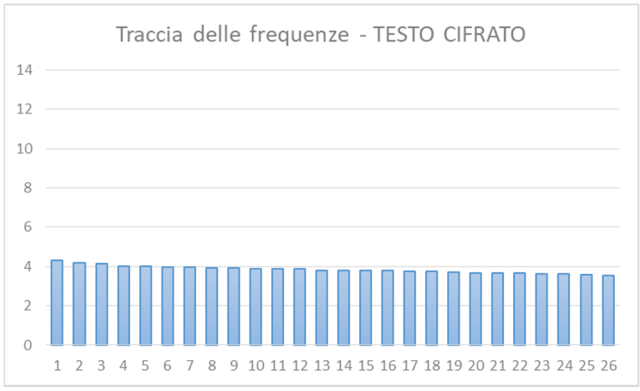
TESTO 3 Lingua:

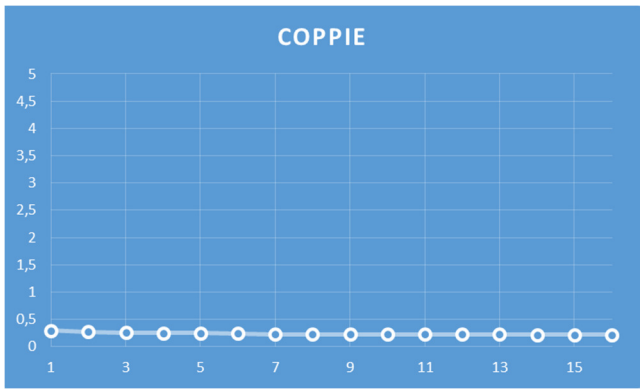


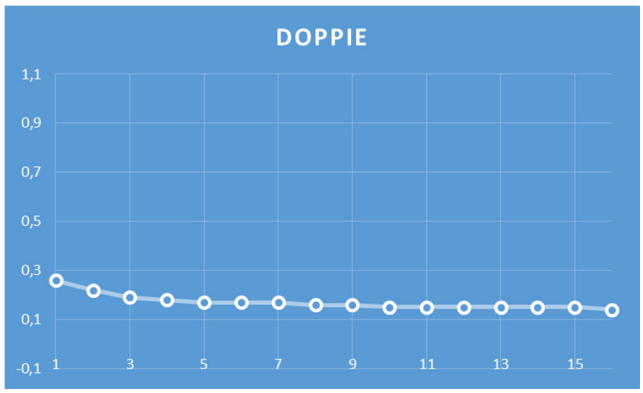




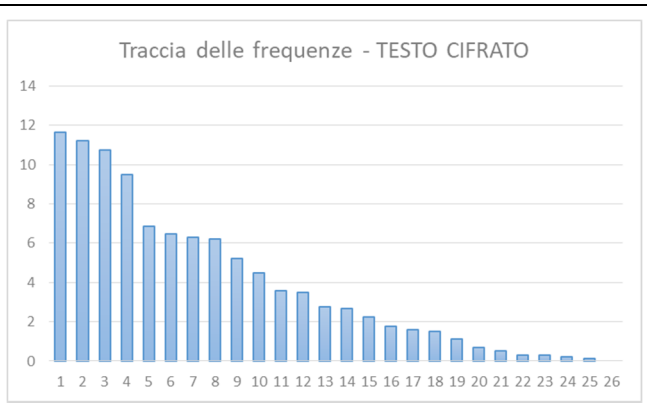
TESTO 4 Lingua:

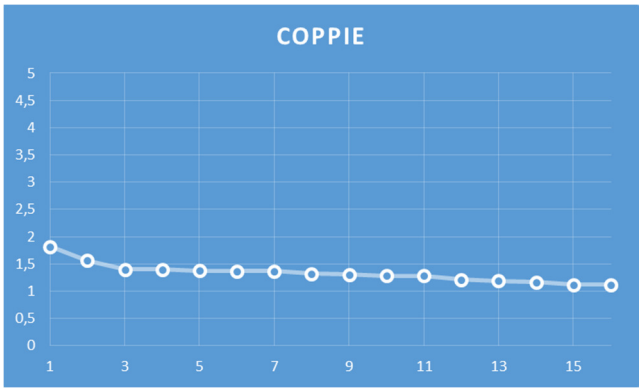


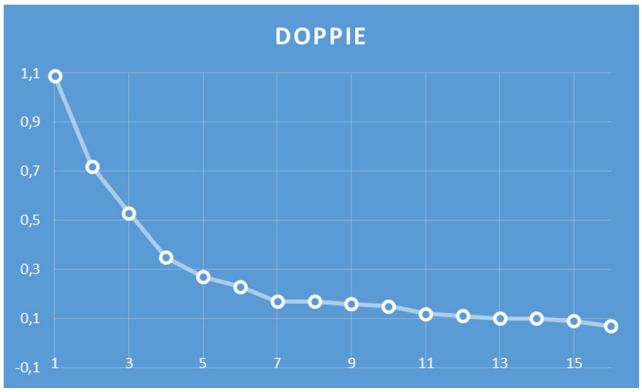




TESTO 5 Lingua:




COPPIE
5
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
1
3
5
7
9
11
13
15


DOPPIE
1,1
0,9
0,7
0,5
0,3
0,1
-0,1
1
3
5
7
9
11
13
15