

Here are some of the most frequent bigrams from the text of selected articles that appeared in *The New York Times* on 17 June 2002. The number after each bigram is the number of times the bigram occurred in about 30,700 letters of text.

TH	909	AT	412	NG	293
HE	749	EN	409	TE	289
IN	655	ST	395	SE	272
ER	547	TI	367	ND	267
RE	452	TO	361	AR	262
ES	452	HA	328	AL	247
ON	432	OR	319	AS	242
AN	427	ED	315	SA	240
NT	416	EA	300	IT	240

Here are some useful heuristics from Helen Fouché Gaines's *Cryptanalysis* book. The idea is to use the bigrams to identify the vowels. The vowels A, E, I, and O are usually among the letters with the highest frequency counts. Furthermore,

1. Letters contacting low-frequency letters are usually vowels.
2. Letters showing a wide variety in their contact-letters are usually vowels.
3. In repeated digrams, one letter is usually a vowel.
4. In reversed digrams, one letter is usually a vowel.
5. Doubled consonants are usually flanked by vowels, and vice versa.
6. Vowels do not often contact one another. Once you identify E, any other high-frequency letter that never touches E is almost certainly O, I, or A.