

ARE ACROSTIC MESSAGES REAL?

A. ROSS ECKLER

Morristown, New Jersey

One of the commonest ways to "prove" that an author is engaging in wordplay is to search for odd patterns of letters in his writings. For example, much is made of the fact that the initial letters of successive lines in a poem spell out a reasonably long word, or that the initial letters of successive lines in a prose passage do the same thing. The question that is almost never put by such discoverers is the following: how likely is it that a word of the same length would have appeared by chance? Only if this probability is very low - say, 0.01 or less - can one confidently assert that the author put the message there.

In particular, let us consider the claim mentioned in the May 1984 *Kickshaws* that Shakespeare deliberately introduced the name TITANIA into the initial letters of a speech uttered by her in "A Midsummer Night's Dream".

Thou shalt remain here, whether thou wilt or no.
I am a spirit of no common rate,
The summer still doth tend upon my state;
And I do love thee. Therefore go with me.
I'll give thee fairies to attend on thee;
And they shall fetch thee jewels from the deep, ..

On the face of it, this appears to be a remarkable coincidence. What is the probability that this happened by accident?

I recorded 1000 consecutive letters beginning the first lines of poetic dialogue in Scene 1 of Act 1, Scenes 1 and 2 of Act 2, and Scene 2 of Act 3 in this Shakespearean play (I omitted Scene 2 of Act 1 and Scene 1 of Act 3 because these are largely written in prose). In this sequence, I noted the appearance of 110 two-letter words, 74 three-letter words, 6 four-letter words (FAST, DATA, WATT, SANE, CHAP, WIND) and 1 five-letter word (CHAPS) from the Merriam-Webster Pocket Dictionary. (Of course, one would not necessarily get 110 two-letter words in a second sample; all one can say is that, with two chances out of three, the number discovered would lie between $110 + \sqrt{110}$ and $100 - \sqrt{110}$, or 99.5 to 120.5.) These data can be adequately summarized by the following empirical rule:

Expected number of words of i letters = $2.2nS/26^i$

where n is the vocabulary size of the dictionary being used, and S is the sequence length. For the Shakespearean data, the predictions generated by this formula are 124, 68, 9 and 1, reasonably

consistent with the observations.

Using this rule, one can predict that in the 100,000 poetic lines of all Shakespearean plays the expected number of seven-letter words from the Pocket Dictionary is $2.2(4591)(100,000)/26^7 = 0.13$; TITANIA looks like a rather unusual event. However, if one allows two letters on the same line as Shakespeare did, then this quantity must be multiplied by six, since one might encounter any of the arrangements TI,T,A,N,I,A; T,I,T,A,N,I,A; T,I,TA,N,I,A; T,I,T,AN,I,A; T,I,T,A,NI,A; T,I,T,A,N,IA. This raises the expected number to 0.78, suggesting that it is not unlikely that a seven-letter Pocket Dictionary word appears somewhere in Shakespeare in a modified acrostical form.

However, we have not yet made use of the fact that TITANIA appears in a speech uttered by Titania. If one asks for the probability that a self-referential acrostic appears somewhere in Shakespeare, the number of allowable words is not 4591 but 1, and the expectation correspondingly shrinks to 0.00016! Accordingly, I conclude that Shakespeare deliberately doctored this passage to spell out Titania's name. Do Shakespearean scholars know of any other self-referential acrostics, or is this one unique?

Actually, the probability of finding Titania in an acrostic is even less, if the question posed is "What is the expected number of self-referential acrostics of actors with seven-letter names?", for there are far fewer than 100,000 lines uttered by such characters. By the same token, the corresponding expectation for a self-referential acrostic of an actor with a (say) four-letter name is much greater, and perhaps one should not be surprised if one is found.

Are the predictions of this rule consistent with the results of the July/August 1980 Games magazine contest to locate accidental literary acrostics; reported on in the August 1981 Word Ways? The winner found the eight-letter word SYNONYMS; no others found words of more than seven letters. Words from Webster's Third Edition were allowed, together with such derived forms as plurals, past tenses and participles; however, it is likely that most readers relied mostly on their word-recognition vocabulary and overlooked the rarer acrostic words. Seventy-seven people sent in entries for the initial contest, which was then extended a short time and the prize doubled from \$150 to \$300. Assuming 100 entrants in all, each one having a recognition vocabulary of 5,000 eight-letter words (there are more than 30,000 eight-letter words in boldface in Webster's Second, not counting derived forms), each entrant would have had to examine a letter-sequence of 190,000 characters for an expectation of one eight-letter word to appear as a result of their joint efforts: $1 = 2.2(100)(5000)(190,000)/26^8$. At the rate of two characters per second, this would take a little more than 26 hours for each contestant to accomplish. (Time spent on checking plausible words in the dictionary is extra.) This seems like a lot of time to spend to win a \$300 prize, but it is certainly well within the capability of a determined and sufficiently optimistic contestant.

Will Shortz reported that in 30 minutes of searching the best he could find was the five-letter word TIGHT. Using the above estimates of recognition vocabulary size and scan rate, the expected number of five-letter words he should have seen is $1.7 = 2.2(5000)$ $(1801)/26^5$.

One can ask whether the 2.2 factor is representative of modern prose. To assess this, I took the initial letters of 1000 consecutive words from Harold J. Leavitt's *Managerial Psychology*, Second Edition (University of Chicago, 1964) and found in this sequence 114, 95, 11 and 1 words of two through five letters. These data suggest a slightly larger value for the multiplicative factor, say 2.5.

QUERY

Louis Phillips of New York City wishes to compile a collection of spelling jokes, and asks Word Ways readers for examples of this genre. Here are three samples:

Teacher: Jimmy, can you spell PHILADELPHIA?

Jimmy: PHILDELPHI.

Teacher: What happened to the A's?

Jimmy: They moved to Kansas City.

Teacher: Jimmy, could you spell WEATHER?

Jimmy: WHIATUR.

Teacher: That's the worst spell of weather we've had in a long time.

Teacher: Jimmy, could you please spell WASP?

Jimmy: WAS.

Teacher: But what's at the end?

Jimmy: Its stinger.

These remind the editor of some related wordplay:

CH CH: what is missing? You are.

ASSUME: something that makes an ass out of you and me.