Ralph G. Beamann
Boothwyn, Pennsylvania

A small group of logologists have long been intrigued by the use of all the alphabetical letters in a single sentence. No, I don't mean such examples as PACK MY BOX WITH FIVE DOZEN LIQUOR JUGS, but an uncontrived and natural use of all the letters.

These are exceedingly rare. In fact, so much so that one delights in finding a sentence having all but one of the letters.

A recent example that caught my eye was the sentence by Lowell Thomas in The Saturday Evening Post for Winter, 1971:

I was in an Assamese tea garden on a very hot day at about three in the afternoon, relaxing after a jungle patrol, when I noticed a quick movement in the bushes across the compound.

This quote misses only the letter Z.

How rare are these? How many letters must one check to find most if not all of the letters?

From data obtained by Bell Telephone Laboratories, I note the percent frequencies of the least common letters are $Z = 0.06$, $Q = 0.10$, $J = 0.10$, $X = 0.16$, and $K = 0.60$. These frequencies are expressed so that the sum of all the letter frequencies is 100.

We see now why Lowell Thomas missed a Z. On the average, it occurs only 6 times in 10,000 letters, or once in 1667 letters. To compound the problem, one must also ensure that all the other letters are included.

The odds for finding all 26 letters in a string of letters selected at random from English text may be calculated by fundamental frequency statistics.

For a 50-50 chance of finding them all, I calculate a Pangrammatic Window at about 2000 letters. Actually, I needed only the letters $Z$, $Q$, $J$, and $X$, for the chance of finding $K$ in a window of 2000 letters is 99.994 per cent. All other letters are even more probable; for example, there should be about 250 E's in a 2000-letter string. The fairly rare letter $X$ is itself only of minor consequence; the three most uncommon letters ($Z$, $Q$ and $J$) suggest a window of 1930 letters.
for an even chance of catching them all.

It's one thing to calculate a mathematical probability, and yet another thing to verify it in practice. Here are the windows I have found for the whole alphabet in recent literature:

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6300</td>
<td>2600</td>
<td>2500</td>
<td>1000</td>
<td>7000</td>
</tr>
<tr>
<td>2600</td>
<td>4000</td>
<td>5250</td>
<td>5500</td>
<td>8000</td>
</tr>
<tr>
<td>4800</td>
<td>2100</td>
<td>750</td>
<td>1850</td>
<td>1500</td>
</tr>
<tr>
<td>2580</td>
<td>750</td>
<td>2500</td>
<td>1100</td>
<td>1500</td>
</tr>
<tr>
<td>2430</td>
<td>7500</td>
<td>1475</td>
<td>1950</td>
<td>8400</td>
</tr>
</tbody>
</table>


This is a very small sample. Will each reader of Word Ways send us her or his own analysis of the True Pangrammatic Window?

The randomness assumptions underlying the calculation of a Pangrammatic Window can be questioned. Some authors are more disposed to using one letter than another (British authors, for example, underuse Z with spellings such as civilization). Some subjects naturally call for more of one letter than another. Even more important, the window width is probably understated because the letter Z is frequently doubled (if all Z's were in pairs in English text, then twice as large a window would be required to have a 50-50 chance of catching a Z). Looking at the median of the 25 observations, I conclude that a pangrammatic window of 2500, not 2000, is required for an even chance of catching all letters.

How narrow can a specific example be? As a quote to beat, I offer the near-pangrammatic excerpt from The Birth of Britain by Winston Churchill, which uses all letters but X in only 84 letters:

... have enjoyed the romantic tragedy of Queen Eleanor
penetrating the protecting maze at Woodstock by ...

I close with a brief note to the mathematically-inclined reader explaining how pangrammatic windows are calculated. Let n denote the number of letters in the window, and p(j) the probability that the jth letter will be drawn at random from English text. The probability that the window will contain one or more examples of the jth letter is equal to \( P(j) = 1 - (1 - p(j))^n \). The nth power can be easily obtained using logarithms, or it can be approximated by \( \exp(-np(j)) \), where \( \exp(a) \) is defined as the numerical constant e (equal to 2.71828...) raised to the ath power. Finally, one obtains the probability that a sample of n letters will contain one or more of each of the letters Z, Q, X and J by the formula \( P = P(z) P(q) P(x) P(j) \). Note that one must obtain by trial and error the n corresponding to a P of 0.50.