WORD WEIGHTS

CHARLES W. BOSTICK
Ashton, Maryland

In his article "Lightweights and Heavyweights" in the November 1972 issue of Word Ways, Darryl Francis introduced the concept of weight for words. He then went on to form two lists, one of words that were heavy for their size and another of words that were light for their size. I'd like to play around with that concept a bit more in this article.

Let each letter, A through Z, have a "weight" that corresponds to its position in the standard English alphabet: A = 1, B = 2, ... , Z = 26. Let the weight of a word be the sum of the weights of its letters. Thus COP has weight 3 + 15 + 16 = 34, and COOP has weight 3 + 15 + 15 + 16 = 49.

With such a definition of weight, it seems reasonable to expect that longer words will generally be heavier than shorter words since there are more letters to sum. In order to compensate for this bias, the weight of a word can be divided by its length (number of letters) to get its "density", a concept which Darryl Francis called "average letter weight". Thus COP has density 34/3 = 11.33, and COOP has density 49/4 = 12.25.

The concepts of weight and density do not depend on letter positions within a word. Thus transpositions such as BAT and TAB have equal weights and densities. If, however, the weight of a letter <u>and</u> its position in a word are considered, a new concept of "center of gravity" can be formulated. It indicates how the weights of letters are balanced in a word, being large if the heavy letters are near the end of the word and small if the heavy letters are near the beginning. To calculate the center of gravity of a word do the following:

(1) write the word	C A T C H
(2) list its letter positions	1 2 3 4 5
(3) list its letter weights	3 1 20 3 8
(4) multiply, position by position, row (2) by row (3)	3 2 60 12 40
(5) add the numbers in row (4)	3 + 2 + 60 + 12 + 40 = 117
(6) add the numbers in row (3)	3 + 1 + 20 + 3 + 8 = 35
(7) divide the sum in (5) by the sum in (6)	117/35 = 3.34

The resulting number is the word's center of gravity. In the case of

CATCH, the center of gravity (abbreviated CG) is almost halfway between the T in position 3 and the C in position 4.

Generally the CG will distinguish between transpositions. Thus, BAT has CG = 2.78 showing the heavy T at the end, but TAB has CG = 1.22 showing the heavy T at the beginning. PRIEST and RESPIT, however, are transpositions which have the same center of gravity (3.61). Can you find others?

Because the notion of center of gravity corresponds to the balancing point of a word, it comes into play when examining centrally balanced beam (CBB) words, discussed by J. A. Lindon in the February 1969 issue of Word Ways. These words are special cases in which the CG is at the exact middle -- the central letter in odd-length words such as SHAFT (CG = 3), or else halfway between the two central letters in even-length words such as AXLE (CG = 2.5). One does not have to go through all the arithmetic to know that all palindromes are centrally balanced beam words.

As with the concept of weights, it seems reasonable to expect that longer words will generally have larger centers of gravity than shorter words. It certainly is true that appending another letter to a word will increase its center of gravity. In order to compensate for this bias, the CG of a word can be divided by the (length + 1) of that word to obtain its "normalized center of gravity" (abbreviated NCG). This measures the proportionate distance along a word where the weights of its letters balance. For example, CATCH has NCG = 3.34/6 = .56, TAB has NCG = 1.22/4 = .30, and SHAFT has NCG = 3/6 = .50. Centrally balanced beam words will automatically have a NCG of one-half.

It's time to take a break. I have defined the four related concepts of weight, density, center of gravity, and normalized center of gravity. Many questions spring to mind:

What is the heaviest word? the lightest?
What is the densest word? the least dense?
What word has largest center of gravity? smallest center of gravity?
What word has largest normalized center of gravity? smallest normalized center of gravity?

The word A clearly is the lightest, the least dense, and has the smallest center of gravity. The reader is invited to try his wits at answering the other questions. As mentioned before, he should look at long words to find heavy words or those with large centers of gravity. Paradoxically, he should examine short words to find those with great density or normalized center of gravity. In Darryl Francis's article, the word densities gradually decreased from 26 to 17 as the word length increased from 1 to 15.

Now consider the possibilities of applying these four concepts to classes of related words. For example, take the twelve months, given in the table on the next page:

	Weight	Density	CG	NCG
January	90	12.86	4.73	.592
February	96	12.00	5.53	.615
March	43	8.60	2.81	.469
April	56	11.20	3.27	.545
May	39	13.00	2.31	.577
June	50	12.50	2.28	.456
July	68	17.00	2.76	.553
August	89	14.83	4.08	.583
September	103	11.44	4.62	.462
October	78	11.14	3.94	.492
November	94	11.75	4.11	.456
December	55	6.88	5.44	.604
Average	71.75	11.93	3.82	.534

The lightest month is May, and the heaviest is September. December is least dense while July is most dense. June has the smallest center of gravity and February has the largest. June also has the smallest normalized center of gravity while February has the largest.

For purposes of comparison, seven classes of words were examined: the 12 months, the first 20 cardinal numbers (one, two, ...), the 50 states, the 104 chemical elements, 66 vegetables, 217 birds and 208 words chosen in a pseudorandom manner from a dictionary (the guide words at the top of every fifth page). The results are given in the tables below, in the form of averages and extremes:

Weight

Months	71.75 (May 39, September 103)
Cardinals	69.30 (one 34, seventeen 109)
Elements	94.57 (lead 22, protactinium 159)
States	93.92 (Alabama 31, Massachusetts 168)
Vegetables	74.32 (cabbage 21, cauliflower 125)
Birds	76.46 (kea 17, yellowthroat 174)
Random	87.83 (each 17, isoagglutination 193)

Density

Months	11.93 (December 6.88, July 17.00)
Cardinals	12.71 (eighteen 9.13, two 19.33)
Elements	12.44 (lead 5.50, yttrium 18.00)
States	11.24 (Alabama 4.43, New York 15.86)
Vegetables	11.48 (cabbage 3.00, soy 19.67)
Birds	11.37 (chickadee 5.44, vulture 17.00)
Random	11.69 (each 4.25, syzygy 21.17)

Center of Gravity

Months	3.82 (June 2.28, February 5.53)

Cardinals	3.24 (one 1.71, eighteen 4.84)
Elements	4.49 (tin 1.86, californium 7.18)
States	4.45 (Utah 1.92, Massachusetts 7.88)
Vegetables	3.85 (yam 1.69, cauliflower 6.85)
Birds	3.90 (kea 1.41, capercaillie 6.88)
Random	4.29 (off 1.67, isoagglutination 8.79)

Normalized Center of Gravity

Months	.534 (June .456, February .615)
Cardinals	.490 (three .402, eight .599)
Elements	.523 (lead .373, cadmium .646)
States	.482 (Nevada .365, Idaho .572)
Vegetables	.514 (rutabaga .308, beans .675)
Birds	.500 (rhea .330, canary .671)
Random	.505 (peace .370, beau .680)

Note that LEAD is the lightest element word. Although ONE is the lightest number word, TWENTY is not the heaviest. Three element words vie for the heaviest with a weight of 159 (I have listed only one). The vegetable word PEA does not occur in the list of 66; the word PEAS is used instead.

I find the average values of each class very interesting. Element words and state words are heavy, the former being quite dense but the latter being least dense. Number words have smaller centers of gravity, and month words tip most to the right.

All these definitions and calculations contribute very little to solving the riddle of the universe. However, they intrigue me and I thought you might like to play with them too.