

WORDS WITH UNUSUAL SHAPES

ALAN FRANK

Medford, Massachusetts

Previous articles in *Word Ways* have discussed such matters as the words with the most hyphens, words with internal capital letters, and various symbols which appear as entries in Webster's Second. These can all be characterized as word shapes. This article will present some of the ways in which shapes can be unusual, including examples and a quiz.

I'll start with the assertion that normal English words consist of a sequence of lower-case letters. The previous sentence illustrates words with the three most common exceptions: apostrophes, capital letters, and hyphens. These will be dealt with in more detail, followed by some stranger shapes. Some words look strange because they were imported from a language with a slightly different alphabet (e.g., résumé). However, these shapes will for the most part not be discussed in this article.

Words with apostrophes arise in three main ways. The two most common are possessives ("the cat's dish") and contractions ("The cat's eating. No, he isn't"). There are a few cases in which a possessive without an apostrophe looks like a plural (its, Pikes Peak), and one case in which a plural with apostrophe looks like a possessive (maitre d's). With respect to multi-word contractions, Lewis Carroll once pointed out that since the apostrophe in negative pronouns is generally taking the place of the letter O in the word not, the contraction for will not should be written wo'n't. But, as usual, language turns up its nose at logic. Other contractions are formed from shortenings of single words. Rarely, the apostrophe replaces the first letter(s), as in 'tis and 'swounds. In some cases the spelling follows pronunciation, such as o' (for of), bo's'n, and fo'c's'le. In old baseball box scores, long names were shortened, as Yastrzemski to Yas'mski; however, today the apostrophes are not so used. The third way in which apostrophes are used is to inflect certain nouns and verbs: x's, OD'ing [6000 Words].

Capitalization in English is used almost exclusively for proper nouns. In most cases, only the first letter is capitalized. In compound names, there may be multiple capitals (FitzGerald, TriBeCa). More unusual cases feature internal capitalization only (terHorst, Tir na nOg). Sometimes, all the letters are capitalized, generally in acronyms (SWAT[2nd Barnhart], SALT[2nd Barnhart]) or words which pretend to be (SNOBOL[6000 Words], HAL[Dickson]). The military delights in jamming together the first few letters of each word of a name to create long acronyms such as COMSERFORSOPACSUBCOM

[Dickson] for "COMmander, SERvice FORce, SOuth PACific SUBordin-ate COMmand". The only other uses I have found for capitalization are individual letters and typhes of logic gates(OR[6000 Words]).

Hyphens are sometimes used to combine two words (or one word and a prefix or suffix) into a compound, such as semi-infinite and hazel-brown. Most hyphenated words have only a single hyphen. Some have more: common sources are phrases turned into adjectives (out-of-the-way) or rarely nouns (thing-in-itself), flowers (jack-in-the-pulpit), and prefixes or suffixes added to words which are already hyphenated (semi--armor-piercing). In the last of these cases, the first hyphen should be an en dash, though such words are rarely printed that way (see "The Long Hyphen" in the May 1977 Word Ways). In what I call conjuncted compounds, a hyphen can appear at the end of a word (twenty- or thirty-fold).

Some dictionary entries are composed of multiple words, such as thunder ax. Depending upon the dictionary, these can get quite long. Alas, computer lists fail at this point, but a cat may look at a king.

The four components of shape introduced in this article - apostrophes (A), capitals (C), hyphens (H), and spaces (S) - appear in every possible combination: Alexander's (AC), will-o'-the-wisp (AH), cat's cradle (AS), Tom-and-Jerryism (CH), New Jersey (CS), warty-faced honey eater (HS), Jack-o'-the-dust (ACH), April Fool's DAY (ACS), rabbit's-foot clover (AHS), Ku-Klux Klan (CHS), and Pons-Winnecke's comet (ACHS). Quadruple combinations are sufficiently rare that it would be of interest to compile a list of Websterian examples.

Dictionaries also provide entries with periods (Mr. Clean[2nd Barnhart], numbers (1068) with subscripts (vitamin B₁), commas(United States Rifle, Model of 1903), diereses (coöperate), and slashes (and/or). Exclamation points are featured in the names of several computer products: Compute!, Friday!, and TK!Solver. This is apparently a different use from the click in !Kung. Chemical names often have weird punctuation; the November 20, 1985 New York Times gave dichloro[4,4-dimethyl-5-[[[(methylamino)carbonyl]oxy]imino]pentanenitrile]-,(T-4)-zinc on the EPA toxic chemicals list.

Once we have all these features, a complete set of examples is no longer possible. Here are selected ones: AC/DC[6000 Words], A-OK[6000 Words], PAVE PAWS[Dickson], RNase[6000 Words], acetyl-coA[6000 Words], blood-CSF barrier[2nd Barnhart], CHiPs[television show], ReGLS[computer language], RaMBaM, A1, M-1[2nd Barnhart], Red No. 2[2nd Barnhart], MacArthur-Forrest process, Riemann-Christoffel three-index symbols, charmed-quark--charmed-antiquark[2nd Barnhart, at orthocharmonium], St.-George's-herb, 'round-the-clock, ne'er-do-well, one-idea'd, tare an' ages, 'n', ha'penny, shuckin' and jivin', and T'n'T[2nd Barnhart, in quotation at Bloody Maria].

As we have seen, words consist of a sequence of lower-case and possibly capital letters, interspersed with hyphens, apostrophes and spaces. It is now time to quantify the concept of shape in

multi-word dictionary entries - that is, look at the word-length patterns that occur. To simplify the discussion, I use National Puzzlers' League notation, in which, for example, Pons-Winnecke's comet is summarized by *4-*8'1 5. Looking only at Webster's Second, most non-solid entries are either singly-hyphenated or two words. All patterns from 1 3 to 11 9 and from 1-3 to 11-9 exist, with 5 4 and 4-6 being the most common unhyphenated and hyphenated patterns, respectively. There are also hundreds of three-word phrases. Here, I note three main categories, of which 4-4 4, 4 4 5, and 3-2-6 are the most common patterns. On the other hand, there are only two dozen words with the pattern *x *y-*z. For some patterns, examples can be found having all possible hyphenations; 4 3 5 yields hard oat grass, rice cut-grass, coal-tar crude and song-and-dance.

A listing of dictionary phrases sorted by pattern shows some examples of what cryptogram constructors have long known: that there are unsuspected constraints on English phrases. For example, of the five 5 3 5 entries, four end in the word grass, and the five 6-4 9 entries all begin with either double or engine. Among the most extreme examples are 3-4 11, which provides dry- and wet-bulb thermometer and 6-3-6, yielding thread-the-needle or needle-and-thread.

For each of the following enumerations, there exists a relatively well-known Websterian entry. How many can you find? (Hint: the answers are in alphabetical order.) If you give up, see Answers and Solutions at the end of this issue.

7-2-3	*5 *7 2 1 *7 *10	5-8-6	3-3 5
3-1'-4-5	*4-*1 4	6 *1	11-3
5 3 5	*9-*5 5	15 7	2-2-3-6
4-1-6-2	6-3 3 (or *9 *3)	3 6 3 6	4-3-3-4-2
*6 *7 7	7-4-3	3-5 8	
8-6 4	3-1'-2-5	9-4 6	
4 5 5 3	10-12	6 15	

With the exception of proper nouns, all untagged words in this article have been taken from the second or third editions of Webster's New International Dictionary of the English Language. Words that are tagged come from:

- [Dickson] Paul Dickson, *Words* (Delacorte Press, New York, 1982)
- [2nd Barnhart] Barnhart, Steinmetz, and Barnhart, *The Second Barnhart Dictionary of New English* (Barnhart/Harper & Row, Bronxville, New York, 1980)
- [6000 Words] 6000 Words, *A Supplement to Webster's Third New International Dictionary* (G. & C. Merriam Company, Springfield, Massachusetts, 1976)