ALPHABETICAL PATTERNS

A. ROSS ECKLER
Morristown, New Jersey

Readers of **Word Ways** should be familiar with the concept of word patterns, exemplified by Jack Levine's three-volume rearrangement of Webster's Second and Third Editions into groups of words having the same pattern, such as EXCESS and BAMBOO. (Word patterns have even been discussed in large-circulation magazines; see physicist Sheldon Glashow's "The Game of Bop" in the Sep/Oct 1992 issue of Quantum.) It is the object of this article to show how the concept of alphabetic patterns can be similarly codified. The alphabetic pattern of a word is ascertained by writing the letters of the alphabet in a row on a slip of paper (in circular fashion, with Z followed by A), and sliding this strip back and forth underneath the word to identify letter-matches (crashes) with the letters of the word. For example, the letters of the word WRETCH match with four different alphabet-shifts, denoted by capital letters in the four alphabets below:

					W	R	E	T	C	H				
W-alphabet	.	.	u	v	W	x	y	z	a	b	c	d	.	.
Q-alphabet	.	.	o	p	q	R	s	T	u	v	w	x	.	.
C-alphabet	.	.	a	b	c	d	E	f	g	H	i	j	.	.
Y-alphabet	.	.	w	x	y	z	a	b	C	d	e	f	.	.

The crashing alphabets are, for convenience, identified by the letter corresponding to the first letter in the word. If a letter matches the A-alphabet, it is said to be invariant.

One can characterize the alphabetic pattern of a word by the identifiers of the crashing alphabets; thus, WRETCH has the alphabetic pattern WQCQYC, with two alphabets (W and Y) crashing once and two more (Q and C) crashing twice. Using this equivalence, one can classify all words in the dictionary by their alphabetic patterns, grouping together those (such as WQCQYC and PNDNXD) whose alphabetic patterns are equivalent. For brevity, each group is identified by its first pattern, alphabetically speaking (ABCBDC). In turn, groups can be assembled into supergroups which are identified by the number of singly-, doubly-, etc. crashing alphabets (WRETCH, for example, has two single-crashes and two double-crashes, as does LANCES, with the different alphabetic pattern LZLZAN; both are in supergroup 2211).

These concepts can be illustrated by classifying the three-letter, four-letter and five-letter boldface words in the Merriam-Webster Pocket Dictio nary according to their alphabetical patterns.

Three-Letter Words (547 words)

Supergroup 111 (498 words)
Supergroup 21 (49 words)
Pattern AAB 12 (deb den dew eft ope opt pro pry sty tub tug tun)
Pattern ABA 21 (arc bad bed bid bud cue egg ilk irk ink moo nap nip par per rat rot rut vex way yea)
Pattern ABB 16 (cab cop dab elm fop gap hop jab lab lop mop nab ode sop tab top)

Four-Letter Words (1954 words)

Supergroup 1111 (1538 words)
Supergroup 211 (392 words)
Pattern AABC 68 (abbe abet able ably abut dead ...)
Pattern ABAC 70 (arch body chew clew crew doff ...)
Pattern ABCA 63 (acid amid arid avid auld awed ...)
Pattern ABBC 46 (alms anon baby chic chid chin ...)
Pattern ABCB 48 (ache acme acne acre ammo away ...)
Pattern ABCC 97 (aide Arab atop balm bast best ...)
Supergroup 31 (14 words)
Pattern AAAB 6 (deft defy nope stub stud stun)
Pattern ABAA 5 (bade bide bode chef clef)
Pattern AABA 2 (abed hick)
Pattern ABBB 1 (erst)
Supergroup 22 (10 words)
Pattern AABB 5 (hide high node stab stop)
Pattern ABAB 2 (grit spur)
Pattern ABBA 3 (babe shiv whiz)
Supergroup 4 (0 words)

Five-Letter Words (2808 words)

Supergroup 11111 (1860 words)
Supergroup 2111 (837 words)
Pattern AABCD 133 (aback abaft abash abbey abbot abeam ...)
Pattern ABACD 80 (ascot badge badly buddy budge cheap ...)
Pattern ABCAD 82 (baker baler bases beset betel bevel ...)
Pattern ABCDA 67 (acute adage adobe addle adore aerie ...)
Pattern ABBCD 68 (adept arson aster astir befit befog ...)
Pattern ABCBD 83 (again along apart aport beige bland ...)
Pattern ABCDB 89 (altho Aztec beach beech belch bench ...)
Pattern ABCCD 97 (adder adopt annoy aphid aught balmy ...)
Pattern ABCDC 63 (bawdy brick brink brisk burnt cache ...)
Pattern ABCDD 75 (angst avast Bantu beano beast blade ...)
Supergroup 311 (46 words)
Pattern AAABC 8 (defer stuck study stuff stump stung stunk stunt)
Pattern ABAAC 6 (bedew bided cleft eight parse purse)
Pattern AABCA 8 (abase abate above abuse death depth straw strew)
Pattern AABAC 4 (deign opera stave stove)
Pattern ABACA 1 (squaw)
Pattern ABCAA 3 (aside brief dough)
Pattern ABBBC 0
Pattern ABBCB 8 (chick chink rabid shirk thick think toper whisk)
Pattern ABCBB 3 (llano neigh weigh)
Pattern ABCCC 5 (burst first verst worst wurst)

Supergroup 221 (61 words)
Pattern AABBC 1 (Norse)
Pattern ABABC 9 (glint grits lance links parch slung slunk spurn spurt)
Pattern ABBAC 2 (babel labor)
Pattern AABCB 3 (noway stick stink)
Pattern ABACB 6 (check cheek retch rotor shuck squat)
Pattern ABBCA 0
Pattern AACBB 4 (deist ghost stoop strop)
Pattern ABCAB 5 (boner borer bower boxer ovary)
Pattern ABCBA 7 (aline alone clang cling clung plant screw)
Pattern ACABB 6 (chest grief grist guide squad sough)
Pattern ACBAB 1 (beret)
Pattern ACBBA 2 (ankle letup)
Pattern ABBCC 8 (chide chief chino etude snoop thief thigh whist)
Pattern ABCBC 4 (byway blink clink slink)
Pattern ABCCB 3 (aphis donor honor)
Supergroup 41 (2 words)
Pattern AABAA 2 (abide abode) (other 4 patterns empty)
Supergroup 32 (1 word)
Pattern ABBAA 1 (anode) (other 9 patterns empty)

There seems no a priori reason to expect any alphabet-shift to be more prevalent than another. This hypothesis can be easily tested by a probabilistic model which assumes that alphabet shifts occur at random and independently with equal probabilities to form words. Thus, for example, the chance that a three-letter word will contain three different alphabetic patterns is equal to (26/26)(25/26)(24/26) = 0.888, or that it will contain two letters corresponding to one alphabetic pattern and the third to a different pattern is equal to 3(26/26)(1/26)(25/26) = 0.111. The table below shows that the expected number of words in each supergroup closely matches the observed numbers.

Supergroup	Obs	Exp
111	498	486
21	49	61
1111	1538	1534
211	392	400
31	14	11
22	10	8

Supergroup	Obs	Exp
11111	1861	1865
2111	837	848
311	46	37
221	61	56
41	2	0.7
32	1	1.4

However, when one looks at the numbers for individual patterns, there is somewhat more variability than the equiprobability model predicts. In particular, there are a larger-than-expected number of words with the alphabetic patterns AABCD and ABCC.

It would be of interest to use a computer to compile a type-collection of supergroups for words of six or more letters. The probability model enables one to decide the likelihood of finding one or more words in a given supergroup, given that a dictionary such as Webster's Second or Third is searched. For example, the Air Force list of Webster's Second (plus a few specialized dictionaries) contains 33,226 eight-letter words, and the probability of a word appearing in Supergroup 2222 is equal to $105(25\cdot 24\cdot 23)/26^7$ =

0.00018; therefore, one may expect (0.00018)(33226) = 6 such words to appear. In fact, WRETCHED, with the alphabetic pattern WQCQYCYW, belongs to this supergroup. As another example, what is the longest word in which each letter has its own unique alphabetic shift (that is, belongs to a supergroup of form 111....1)? The table below suggests that a 17-letter word (maybe even an 18-letter one?) can be found.

Word Length	Dictionary Size	Probability	Expected Number
12	23999	.0485	1164
13	17723	.0261	463
14	11910	.0131	156
15	7438	.00602	45
16	4380	.00255	11
17	2504	.00098	2.5
18	1289	.00034	0.4

It should be a straightforward matter to program a computer to discover how many words with the requisite alphabetical patterns appear.

Certain aspects of alphabetic patterns have been touched on in earlier **Word Ways**; the subject was introduced in a Query in May 1972. Invariant letters ending words (A, oB, saC, leaD, statE, etc.) were explored by Darryl Francis in May 1971. Ed Wolpow listed words having four crashes with a particular alphabet in February 1979. The words iNOPeRaTiVe and coOPeRaTiVelY both have six crashes with a single alphabet. Shiftwords (such as JOLLY to CHEER, or OHM to PIN) always have the same alphabetic pattern.

Alphabetical patterns are analogous to the word patterns alluded to at the beginning of this article; in fact, any question about one can be rephrased as a question about the other. The key is to visualize the alphabetic pattern as a "word" in its own right. Thus, the search for the longest word in which each letter has its own unique alphabetic shift (discussed above) corresponds to the search for the longest isogram (a word with no repeated letters), which happens to the the 15-letter word DERMATOGLYPHICS. Similarly, the search for the word with the most letters matching a single shift-alphabet translates to the search for the word with the most repetitions of a single letter, which happens to be hUmUhUmUnUkUnUkUapUaa (among others) with nine. And what is the pattern word corresponding to the longest word with each shift-alphabet repeated twice (WRETCHED, noted earlier)? This is, in fact, a pair isogram; the longest example is SCINTILLESCENT. Words such as undeRSTUdy or liMNOPhile have four consecutive letters in the same shift-alphabet; however, there are no Websterian words with a pattern of more than three consecutive identical letters (waLLLess, headmistreSSShip). No doubt **Word Ways** readers can discover other analogies.

Chris Cole has kindly supplied the author with a list of boldface uncapitalized unhyphenated words with six letters matching a single shift-alphabet. All can be found in one or more of three sources: the Second and Third Editions of the Merriam-Webster Unabridged, and the Random House Unabridged. By far the commonest match is OPRTVY, illustrated by coOPeRaTiVelY, intraOPeRaTiVelY, OPeRaTiVelY, OPeRaTiVitY, preOPeRaTiVelY, postOPeRaTiVelY, and uncoOPeRaTiVelY. NOPRTV is found in iNOPeRaTiVe, iNOPeRaTiVeness, and noNOPeRaTiVe; NOPTUY, in iNOPporTUnelY and iNOPporTUnity; and NPRTUY, in noNsPiRiTUallY and uNsPiRiTUallY. Other matches are illustrated by DauGHterLiNesseS, DEFinItiveNesseS, gyMNOPlaST, nEiGHborLiNesseS, unDEFendabLeNesseS, baLaNOPlaSTy, and NOnfeSTiVelY.

Here is the corresponding list of words with five letters matching the A-alphabet: ABuDEFduf, AgammaGlobuLiNemiaS, ArChEncephaLoN, ArChEtypIcaL(ly), syngEnesIotraNsPlanTation, ABaDEnGo, AntiantHropoMorPhiSm, ApoDEictIcaL(ly), ArChErsHIp, nonDEFeasibLeNess, nonDEFensibLeNess, nonDEFinIteLy, and nonDEFinItiveNess.

NOT ANOTHER PALINDROME BOOK?

Yes, another ... but Go Hang a Salami, I'm a Lasagna Hog (Farrar, Straus, Giroux, 1991; $12.21) is certainly worth adding to the canon. John Agee's cartoons of the 56 palindromes are generally clever, even when the original palindrome seems unpromising ("No, son" is uttered by a father reading a newspaper as his scion, hiding behind the TV, prepares to light a fuse leading to a box of TNT on which his sister is tied). Others, of course, cry out for an apposite cartoon ("Not a banana baton" or "Pooh's hoop"). There's even one Cheater's Palindrome: a car, a man, a Maraca. Some care has been taken with the arrangement of the pages ("Yell alley" opposite "Yawn way", and "Dump mud" opposite "Emil's niece, in slime"). This book deserves a sequel, preferably before the year 2002.