# Word Groups With Mathematical Structure

DR. A. ROSS ECKLER
*Morristown, New Jersey*

Logophiles have extensively studied interesting properties of isolated words: palindromes, anagrams and antigrams, multiple transposals, words with successive doubled letters, and so forth. Somewhat less effort has been expended upon an equally interesting and potentially far richer field: the construction of groups of words which conform to certain rules. Perhaps the best-known examples of these are word ladders (less, loss, lose, lore, more) and word squares.

In this article, I shall describe a number of less well-known but equally challenging word games. The common thread connecting them is mathematical; in all cases, there exist analogous patterns studied and used by algebraists, cryptographers, statisticians, and mathemagicians. The rules of the games all have the following general form:

(1) Select a subset t of the letters in the alphabet;
(2) Generate r copies of each letter;
(3) Arrange these tr letters in groups according to a specified rule: and
(4) Transpose the letters within each group to form an English word.

*Balanced Incomplete Block Designs*

Using t different letters from the alphabet, each one exactly r times, construct b words, each word containing k different letters (obviously, tr is equal to bk). The difficulty of this task is increased by an additional requirement: each pair of different letters must appear together in precisely j of the b words. It is not hard to show that j is equal to r(k—1)/(t—1).

Let us illustrate this with a simple example. Consider the words YEA, PER, YET, PAY, PRY, TAP, ARE, RAT, TRY, and PET. The six letters A, E, P, T, R, and Y have each been taken five times to form a stockpile of thirty letters; these have been arranged in ten words containing three different letters each. Note that the letters Y and T appear in the same word twice (YET, TRY), that the letters P and E appear in the same word twice (PET, PER), and so on for all fifteen two-letter combinations. As an aid in classifying these designs, we use the number-triple (t, b, k); thus, the above words illustrate the (6, 10, 3) design.

Balanced incomplete block designs can be divided into two classes—symmetric designs in which b is equal to t, and asymmetric designs in which b is greater than t. It appears quite difficult to construct word groups for the asymmetric designs; besides the (6, 10, 3) design given above, I know of only one other: (9, 12, 3). A group of words satisfying this design is EMU, THY, AGO, TAU, GYM, HOE, YOU, HAM, GET, HUG, TOM, and YEA. Note that, in this design, each letter-pair appears in only one word.

I suggest several of the simpler asymmetric designs which might yield to logological research. The necessary letter groups in the various words are presented in lower-case type; the reader must replace these letters with his desired choice of t different letters, and then rearrange the letters to form words.

( 8, 14, 4) abcd, abgh, acfh, adfg, efgh, cdef, bdeg, bceh, abef, aceg, adeh, cdgh, bdfh, bcfg
( 9, 12, 6) defghi, abcghi, abcdef, bcefhi, acdfgi, abdegh, bcdegi, acefgh, abdfhi, bcdfgh, acdehi, abefgi
( 6, 15, 4) cdef, bdef, bcef, bcdf, bcde, adef, acef, acdf, acde, abef, abdf, abcf, abce, abcd, abde
( 5, 10, 3) abc, abd, abe, acd, ace, ade, bcd, bce, bde, cde
(10, 15, 4) bcej, aefg, dehi, abeh, bcdf, bfgh, fhij, acfi, cdgh, cegi, bgij, adgj, abdi, defj, achj
(10, 15, 6) adfghi, bcdhij, abcfgj, cdfgij, aeghij, acdeij, abcdeg, bdeghj, abefij, abdfbj, acdefh, bcefhi, cefghj, abcghi, bdefgi

The last two designs are said to be complementary; the letters in corresponding words from the two designs exhaust the original alphabet and do not overlap each other. The simplest designs using five-letter words are appreciably larger: (10, 18, 5) and (9, 18, 5).

Let us now turn to symmetric balanced incomplete block designs (in which b is equal to t). It is possible to obtain solutions for several:

(3, 3, 2) BE, BY, YE
(4, 4, 3) SET, SEA, SAT, EAT
(5, 5, 4) NEAT, SANE, NEST, TANS, SEAT
(6, 6, 5) RATES, CASTE, CREST, CARTS, CARES, CRATE
(7, 7, 6) SPLINT, PLANTS, PLIANT, PAINTS, PLAINS, INSTAL, PLAITS
(7, 7, 3) ADO, ORE, BAR, BOY. YEA, BED, DRY
(7, 7, 4) NAIL, SALE, SINE, LEND, IDEA, SAND, SLID
(8, 8, 7) STINGER, GAITERS, RETAINS, SEATING, STRANGE, RATINGS, GRANITE, ERASING
(9, 9, 8) TRINODES, NOTARIES, INTRADOS, TORNADES, ASTEROID, STRAINED, SEDATION, RATIONED, DONARIES
(10, 10, 9) STERCOLIN, RELATIONS, CONTRAILS, CONSERTAL, CREATIONS, SECTORIAL, LARCENIST, SECTIONAL, CROTALINE, CENSORIAL
(11, 11, 10) PROCLINEST, INTERPOSAL, PROLACTINS, NECROPLAST, PREACTIONS, CARPOLITES, INTERCLASP, CLARIONETS, PLEONASTIC, PRATINCOLE, PSILOCERAN

In the last of these designs, PROCLINEST is the solemn or poetic second person singular form of the verb "to procline," meaning "to lean forward," and NECROPLAST is a word taken from a scientific dictionary and definable as "a mass of dead protoplasm." An alternative choice is the coined word NARCOLEPTS ("persons seized with narcolepsy").

Again, I suggest several additional designs which might yield word groups:

(11, 11, 5) aefgi, bfghj, cghik, adhij, beijk, acfjk, abdgk, abceh, bcdfi, cdegj, defhk
(11, 11, 6) bcdhjk, acdeik, abdefj, bcefgk, acdfgh, bdeghi, cefhij, dfgijk, aeghjk, abfhik, abcgij
(13, 13, 7) abcj, defj, ghij, adgk, behk, cfik, afhl, bdil, cegl, aeim, bfgm, cdhm, jklm

In *BEYOND LANGUAGE* (Charles Scribner's Sons, New York, 1967), Borgmann introduces the (7, 7, 3) and (13, 13, 7) balanced incomplete block designs in Problem 122 (Finite Projective Geometries). These two designs have the properties that (1) each pair of different letters appears together in precisely one word (j equal to unity), and (2) any two words have only one letter in common (a restriction not present in most balanced incomplete block designs). Borgmann proposes the not entirely satisfactory solution AIRY, AUTO, DIEU, KINO, OYES, PANE, PITS, PROD, PUKY, RUNS, SKAD, TREK, and TYND for the latter design (TYND is a variant of "tine" used only in northern England and in Scotland).

Balanced incomplete block designs are extensively used by statisticians in designing field experiments. Their principal value lies in the fact that a large number of different treatments (alphabetic letters) can be systematically applied to blocks of limited size (words of limited length). For example, the statistician may wish to evaluate the comparative nutritive value of twenty different mixes of pig feed, but is limited to pig litters of a maximum size of twelve.

No balanced incomplete block designs exist for b less than t. However, the logologist need not be deterred by this fact; he can create additional word designs having slightly different properties. For each asymmetric balanced incomplete block design (t, b, k) one can define a dual design (b, t, bk/t)—in other words, the roles of b and t, and of k and r, have been interchanged. These designs no longer have the property that each pair of letters appears together in j different words, but instead have the property that any set of m words has exactly one letter in common.

These designs appear to be somewhat easier to solve. Some examples:

( 6, 4, 3) CAN, COT, ATE, ONE
(10, 5, 4) SCAN, SORE, COIL, MARL, MINE
(15, 6, 5) WRIST, WHOLE, CHARM, COUNT, PLAIN, SPUME
(10, 5, 6) GRINDS, MATING, GRATES, REMAND, MISTED
(10, 6, 5) OGLED, GRAPE, POISE, GRIDS, PLAID, SOLAR
(12, 9, 4) SLAM, MORE, MIND, NEAT, TOLD, STIR, BARD, SNOB, BILE

The last two designs can be recognized as duals to the only two asymmetric balanced incomplete block designs for which words have been found. The following design appears to be considerably more difficult:

(14, 8, 7) abcdijk, abghimn, acfhjln, adfgklm, efghijk, cdefimn, bdegjln, bcehklm

*Word Progressions*

Select t different letters from the alphabet and arrange them on the perimeter of a circle. For each of the t groups of k adjacent letters, transpose the letters to form an English word. This problem becomes somewhat more interesting if posed in a slightly different form: for various values of k (2, 3, 4, 5, etc.), what is the largest value of t for which this can be accomplished?

Readers of Borgmann's *BEYOND LANGUAGE* will immediately recognize Problem 26 (A Magic Circle) as a special case of this problem in which t is equal to eight and k is equal to four. It is also clear that when t is one greater than k, a solution of the original problem is provided by a symmetric incomplete block design of the form (i, i, i — 1). For examples from (3, 3, 2) to (11, 11, 10), refer to the preceding section.

I do not know the best possible solution to the word progression problem; readers may be interested in seeing how far they can surpass the solutions below.

(a) For k = 2, O S U P I N E B Y M A D (12 letters)
SO, US, UP, PI, IN, EN, BE, BY, MY, AM, AD, DO
(b) For k = 3, B A J W E F L Y P S U N R O D H I T (18 letters)
JAB, JAW, JEW, FEW, ELF, FLY, PLY, SPY, SUP, SUN, RUN, NOR, ROD, HOD, HID, HIT, BIT, BAT
(c) For k = 4, R B A K C P E O D L G I T S U N (16 letters)
BARK, BACK, PACK, PECK, COPE, DOPE, LODE, GOLD, GILD, GILT, GIST, SUIT, NUTS, RUNS, BURN, BARN
(d) For k = 5, S L O R A C H T E N I W D (13 letters)
SOLAR, CORAL, ROACH, CHART, TEACH, TENCH, THINE, TWINE, WINED, WINDS, WILDS, WOLDS, LORDS
(e) For k = 6, E N D I R A S O L M (10 letters)
RAINED, DRAINS, RADIOS, SAILOR, MOLARS, AMOLES, MELONS, DOLMEN, LIMNED, REMIND

In order to exceed the above values of t for k equal to 2 and 3, it will be necessary to use words consisting entirely of consonants (SH and CWM, for instance).

*Code Word Construction*

When constructing books of code words, cryptographers usually want the separate words to be as different as possible in order to avoid confusion caused by errors in transmission. Thus, the words BENCH and BUNCH are a bad pair because if an error is made in the second letter the recipient cannot be certain which code word was originally sent. However, if the words RIGHT and GRAZE are used, two errors can be made in either word and it can still be unambiguously decoded. (Under some circumstances, even four errors won't interfere with the decoding of these two words.)

These considerations are reflected in the rules governing the following word game. Select t different letters of the alphabet and arrange them in a vertical column to form the initial letters of t words. Rearrange the same t letters in a column to the right, forming the second letters of each word. Continue this process until t k-letter words have been formed. Again, this problem can be recast in a more challenging form: for a given value of k, what is the maximum possible value of t?

In *BEYOND LANGUAGE,* Borgmann introduces a related game in Problem 86 (Irrelevance). However, he relaxes the requirement that each vertical column must consist of the same t alphabetic letters—auy t different ones will do.

This problem is somewhat more difficult than the ones discussed earlier. Readers are invited to improve on the following solutions:

For k = 2: AS, SO, MA, OR, RE, EM (6 words)
For k = 3: YES, RYE, PRY, ASP, SPA, EAR (6 words)
For k = 4: (a) TYRO, YEAR, ESPY, AREA, STOP, POST, OPTS, RAYE
(b) RANI, ONTO, ARID, IDEA, DIRT, NODE, TEAR, ETON

Neither four-letter solution is entirely satisfactory, as each requires one proper name (RAYE, ETON). I have no solution of any sort for words of five or more letters. Note that the solution permits repetitions of the same letter in a single word.

*A Card Trick Mnemonic*

In a well-known card trick, the mathemagician deals ten pairs of playing cards face down and invites the subject to look at the face values of one pair while the mathemagician's back is turned. The mathemagician then deals the cards out in a four-by-five array with card faces upwards. When the subject is asked to say in which of the four rows his two cards appear, the mathemagician immediately identifies the two cards.

The successful execution of this trick depends upon the fact that there are exactly ten different ways in which a pair of cards can be distributed among four rows: both cards in row 1, cards in rows 1 and 2, cards in rows 1 and 3, . . ., both cards in row 4. The mathemagician uses the mnemonic BIBLE, ATLAS, GOOSE, and THIGH to aid in placing the cards in the four rows: the first pair of cards occupy the positions of the B's in the first word, the second pair of cards occupy the positions of the I's in the first and fourth words, and so on.

Logologists can at once generalize this problem. Select k(k—1)/2 different letters from the alphabet; using pairs of these letters, construct (k—1) words of k letters each, with the property that each letter-pair can be uniquely identified from a knowledge of the words (or word) in which the letter-pair appears. For k equal to three or four, the problem is trivial; for k equal to five, one possible solution has already been given. (How many others are there?) For k equal to six, I suggest the words LIVELY, RHYTHM, MUFFIN, SUPPER, and SAVANT. Logologists are invited to try their hand at finding a solution when k is equal to seven, a feat which requires using 21 of the 26 alphabetic letters.