

# LEVINE'S ISOMORPH DICTIONARY

A. ROSS ECKLER

Morristown, New Jersey

Among word buffs, 1971 will undoubtedly be remembered as the year that the Compact Edition of the Oxford English Dictionary was published, making this monumental work available at one-third the price and one-sixth the bulk of the original. An exceedingly useful lexicographic tool has been placed in the hands of many who formerly had to make a trip to the library to consult it. By contrast, one of the least-heralded publishing events of 1971 was the appearance of Jack Levine's A List of Pattern Words of Lengths Two Through Nine. Nevertheless, I predict that the Levine dictionary may have a greater impact than the COED on word buffs. The information in the COED has been available in the OED for decades, but Levine's dictionary enables the logologist to view Webster's Unabridged in an entirely new light: specifically, it groups together all words with the same underlying pattern, such as EXCESS and BAMBOO (and, in fact, 23 rarer words also having the letter-pattern abcadd). Furthermore, the COED costs \$75, but the Levine dictionary may be obtained free while the limited supply lasts (Box 5548, State College Station, Raleigh, North Carolina 27607).

A more careful description of the Levine dictionary may enable readers to assess its logological potential. It consists of nearly 400 pages of computer printout, reduced in size to 8 1/2 by 11 inches. Approximately 184,000 different entries are listed -- boldface single-word entries from Webster's Unabridged (both the Second and Third Editions), and derived words such as plurals, participles and past tenses of verbs. The dictionary lists words in order of increasing length. Words of a given length are partitioned into supergroups defined by the number of letters appearing doubly, triply, quadruply, etc. in them. For example, the word FLORISTS, with one letter (S) repeated twice, is indexed by the partition code 2, and the word HOMOMORPH, with two letters (H, M) repeated twice and one letter (O) repeated three times, is indexed by the partition code 223. Finally, words of a given length in a given supergroup are further divided into groups according to their underlying pattern. For example, the words HORROR and EFFETE both have the partition code 23, but HORROR is defined by the numbers 2.5-3.4.6 (the O appears in the

second and fifth positions, and the R in the third, fourth and sixth positions), whereas EFFETE is defined by the numbers 2.3-1.4.6. This classification system is much more complicated to describe than to use; one quickly learns how to locate any given word-pattern in the dictionary. (In my November, 1971 article calling for an isomorph dictionary, I called the supergroups anagrammatic isomorphs, and the groups isomorphs.)

The Levine dictionary is the first volume of a proposed trilogy; two additional volumes will be issued in 1972 covering words of ten through sixteen letters. Together the three volumes will contain about half a million words. The reader should be warned that none of these volumes list isograms (also called nonpattern words). However, Jack Levine has already listed such words in A List of Words Containing No Repeated Letters, published in 1957 and now out of print.

Levine's original purpose in compiling this dictionary was for aid in solving cryptograms of an elementary nature -- substitution ciphers in which the divisions between words are preserved. For some substitution ciphers, it appears to be a powerful tool indeed. Consider the following cryptogram submitted by Mary Youngquist to the October 1971 Enigma (The National Puzzlers' League):

MEANWITCHIN NDQHDC CITVVZCIV VEQM MEN, NEVQZVVIV  
VSOKHDOV, VZPPIVHV VJEBH HDAVEUUIQHDOS.

No need to indulge in laborious computations of letter-frequencies and adjacent-letter-pair frequencies (statistics which are often distorted by the clever cryptographer). The eye is immediately drawn to CITVVZCIV, a nine-letter word unusually rich in repeated letters. Entering Levine's dictionary with the partition code 223 and the pattern 1.7-2.8-4.5.9, one finds a single word meeting these specifications: REASSURES. Turning to NEVQZVVIV (code 4, pattern 3.6.7.9) one finds only three possible words: DISCUSSES, ABSCISSES, and POLYALLEL. Since V has already been identified as plaintext s, I as plaintext e, and Z as plaintext u, it is obvious that DISCUSSES is the desired word. Turning next to VZPPIVHV (code 23, pattern 3.4-1.6.8), one is a bit dismayed to find eleven words, but one can conclude that the word must be of the plaintext form su--es-s, and the only candidate is SUGGESTS. The fourth usable word is VSOKHDOV (code 22, pattern 1.8-3.7), for which there are listed 115 words. However, the plaintext form must be s---t--s, and once again the only word satisfying these constraints and adding new letters to our stockpile is SYMPTOMS. By now we have identified 14 of the 20 different letters in the cryptogram, and it is a simple matter to guess the rest and learn:

KINDHEARTED DOCTOR REASSURES SICK KID, DISCUSSES SYMPTOMS, SUGGESTS SWIFT TONSILLECTOMY.

Had the longer Levine dictionaries been available, the logical entry-point to the cryptogram would have been HDAVEUUIQHDOS.

At this point, some people will no doubt bewail the fact that the Levine dictionary has taken all the fun out of cryptography by making it a routine dictionary look-up. But no tears need be shed over the imminent demise of the American Cryptogram Association; many cryptograms escape Levine unscathed. All one has to do is construct a cryptogram for which every word has a dozen (or a hundred) alternative patterns; the number of possibilities that one must juggle quickly gets out of hand. In general, the cryptographer should lean toward relatively short words with as few repeated letters as possible.

Word Ways readers are probably more interested in the impact of the Levine dictionary on logological research. It sheds new light on a wide variety of fields; its power can best be illustrated by indicating the modifications in earlier Word Ways articles had this work been available.

1. To begin with, what letter-patterns were overlooked in the sample of an isomorph dictionary in the November 1971 Word Ways? Examples for one-third of the missing four-letter words are found in Levine:

abcb TYPY	abac TYTE
abcc TYPP, SUAA, AIEE	aabc LLYN, LLEW, LLEU
abca WYLW	aabb OAAA
abbc YNNS, POOA	

Several new six-letter words were uncovered, too:

aabcad AASBAI	aabcba CCAPAC	ababab AIAIAI
aabxda EESOME	caabba KOOSSO	abccab SHOOSH
aacbbd AALIIS	acbaab CRACCA	
cbabaa DANANN	babbab KAKKAK	

2. The isomorph concept was explicitly introduced in the August 1969 Kickshaws when Dave Silverman asked readers to find two isomorphs to SWEETHEART and BLOODHOUND. The full answer to this must await the next Levine volume, but his shorter examples can be reexamined. He claimed that ALFALFA and ENTENTE were a unique pair, but Murray Pearce pointed out in the August 1971 issue that SEMSEMS, SIMSIMS, SOUSOUS and HATHATH also exist. Levine's dictionary adds SAM\$AMS and SARSARS to

the collection, but inexplicably omits Murray Pearce's last two words (below the line in Webster's Second). Walter Penney and Murray Pearce found a number of isomorphs of ROCOCO and ILLICIT, but again Levine had the last word, listing five new isomorphs of the former and eleven new isomorphs of the latter. No Word Ways reader contributed an isomorph to NONSENSE, but according to Levine there is one: PAPIOPIO.

3. In the August 1971 Word Ways, Murray Pearce presented, for each letter of the alphabet, the longest known palindrome centered on that letter. A search of the Levine dictionary produced longer palindromes for only two central letters: SOHOS (if the non-Websterian word NARIHIRAN is ignored) and KUMUK. In the same article Murray Pearce also presented, for each letter of the alphabet, the longest known Heads 'n Tails word centered on that letter. Here the Levine dictionary was of considerably more help, replacing two five-letter words and four seven-letter words with six nine-letter words: MUCKAMUCK, OKEYDOKEY, TRINITRIN, OVERLOVER, TARANTARA and TOOKYTOOK. On the other hand, Levine missed the seven-letter Heads 'n Tails word CHERCHE.
  
4. The Switch Puzzle was introduced to recreational mathematics by Sam Loyd (Puzzle 54, Mathematical Puzzles of Sam Loyd, Dover, 1959) and later presented by Henry Ernest Dudeney (Puzzle 293, 300 Best Word Puzzles, Scribner's, 1968). Visualize a horizontal box just large enough to contain twelve square tiles in a line, and a vertical box intersecting the horizontal box at the position of the tenth tile. The vertical box has room for exactly nine tiles above the intersection and two below the intersection; thus it, too, can hold all twelve tiles. If the letters of the word INTERPRETING are inscribed on the tiles, the tiles can be moved one at a time to the vertical box where the same word will appear in columnar form (a "move" is defined as the sliding of one letter any distance within either box, whether or not a corner is turned). INTERPRETING is the only known twelve-letter word that can be so transferred in exactly twelve moves; other words require the solver to slide one (or more) tiles at two separate times. For further details, the reader is referred to the February 1968 and February 1969 Word Ways, where the general patterns of switch words are defined, and switch words suitable for boxes of shorter lengths and various intersection-points are exhibited. Switch words form a broad class of pattern words which include palindromes, reduplications, and some Heads 'n Tails words as special cases. The February 1969 article gives examples for all 8 of the distinct five-letter switch words, 7 of the 11 distinct six-letter switch words, 12 of the 17 distinct seven-letter switch words, and 11 of the 20 distinct

eight-letter switch words. A quick search of Levine revealed a few overlooked patterns:

abccab SHOOSH	abccabd RATTRAP	abccddcab TEAMMATE
bcbcaa SUSURR	abccddbc PIERRIE	abccabdd GREEGREE*
	ababdc TOTOABA	abbacdd TEETERER*

Note that the two asterisked words are defective -- they do not have the most general possible pattern.

5. The usefulness of the Levine dictionary is not restricted to words with completely specified patterns. In the November 1969 Word Ways, Faith Eckler looked for unhyphenated words in Webster's Second Edition having three consecutive identical letters (no hyphens allowed). Only four words were found in a subsequent search of a magnetic tape containing Webster's Second: WALLLESS, GODDESSSHIP, HEADMISTRESSSHIP and PATRON-ESSSHIP. A search of all partition codes containing a 3 (or larger number) turned up the additional words BITTTIEST, SHOOOSHES, SLOOOSHES, HALAAA and VEEED -- but none of these was confirmed by a search of Webster's Second and Third. Equally puzzling was the listing of BOSSSHIP in Levine, which appears as a hyphenated word in Webster's Second. If BOSS-SHIP is admitted by Levine, why are the hyphenated entries GALL-LESS, WALL-LIKE and CROSS-STAR (among many others) omitted?
6. Darryl Francis spent two years searching for a fourteen-letter pair isogram (a word in which each letter appears exactly twice) before coming up with TAENIODONTIDAE for the August 1971 Word Ways. (In the same issue, Ralph Beaman presented CICADELLIDAE and HAPPENCHANCE.) Although the present Levine volume sheds no light on this research, it is clear that it will be a very simple matter to check the next two volumes and find out whether any more long pair isograms exist. Two other questions, however, can be settled immediately: there are no nine-letter trio isograms, and KUKUKUKU is the only eight-letter quadruple isogram.
7. The problem of finding the shortest words containing exactly  $n$  of a given letter of the alphabet was introduced (for  $n = 3$ ) by Dmitri Borgmann on page 153 of Language on Vacation (Scribner's, 1965). He followed this up with an article (for  $n = 4$ ) in the March 1969 Word Ways, and Darryl Francis extended the research (to  $n = 5$ ) in the November 1970 Word Ways. Although Levine's present volume is less useful than the next two will be, it does suggest various improvements for all three lists. Dmitri's first list can be improved by the words DODD, FAFF, GEGG and TETT;

on the other hand, Levine omitted KAKKE and VIVA-VOCE. Dmitri's second list can be improved by the words ALABAMA, DODDED, INIMITIE, KAKKAK, PALLALL, NUNNING and TIT-BITTY; Darryl's list can be improved by the words ANAPANAPA, BEEKEEPER, NONTANNIN and TITTATTOE.

8. In the article "Word Groups With Mathematical Structure" in the November 1968 Word Ways, the following problem was proposed: from a doubled set of 21 different letters, construct six seven-letter words in such a way that (1) each word contains five different single letters and one doubled letter, and (2) each pair of words contains exactly one letter in common. (Such a group of words is useful as a mnemonic in performing various card tricks; magicians will quickly recognize the simpler set BIBLE ATLAS GOOSE and THIGH.) In Levine's dictionary, all such seven-letter words have a partition code of 2 and are grouped together, making it easy to search for suitable candidates. If Y is counted as a vowel, the typical word can have only two vowels; any word with three must be balanced by another with only one. Therefore, it seemed reasonable to search through Levine and collect (1) one-vowel words, and (2) two-vowel words with one vowel Y; careful selection of one word in the first category and two words in the second category should go a long way toward solving the problem. In a first attempt, THRIFTS DYBBUKS and GLOWFLY used up many rare letters, and two more words could be added in many ways, but the sixth word was tantalizingly elusive. In a second attempt, it was noted that RHYTHMS was the only word in Levine satisfying both conditions, and that it fitted well with KNUCKLY, KNUBBLY or KNOBBLY. The rare-letter words TWIZZLE and GUFFAWS quickly led to the solution given above, which uses words given in boldface in Webster's Third.

MEACOCK  
RODDING  
GUFFAWS  
TWIZZLE  
RHYTHMS  
KNUBBLY

Some readers of Word Ways will maintain that Levine's dictionary has taken all the fun out of word-searching in Webster's. I prefer to take a more positive outlook. No doubt certain pastimes (such as searching for pair isograms) are now obsolete, but I believe that far more new doors of logological investigation have been opened. I look forward to many new articles in the pages of Word Ways exploiting these opportunities.