

# TEN-SQUARE HUNTERS, UNITE!

TED CLARKE

Newquay, Cornwall, England

According to the wordsquare pundits, it would appear that they give themselves Buckley's of creating a generally-acceptable ten-square. Nevertheless, although there is widespread agreement among them that the time required to exhaust all possibilities could well be extra-universal, many are devoting Sisyphean time and effort to the compilation of ever-lengthening lists of ten-letter words.

Chris Long's computer estimate in the February 1993 *Word Ways* suggests that the minimum support for a ten-square is around 250,000 words. There is a question as to its validity, prompted by the doubtful application of rigorous mathematical analysis to the capricious nature of word structure. Recent articles, in fact, have displayed wide variation in the empirically-determined support.

Whereas a list of 50,000 or so ten-letter words and phrases can be extracted from major standard dictionaries, formists are struggling to create databases of more than double this size. Where are the extra words to be found? It is probable that there are millions of words from which the ten-letter set could be gathered; there are, for example, more than a million recognized species of insects. Purists, however, would not allow them as valid ten-square contenders. They would also reject all other proper nouns, i.e., names of beings and things normally spelled with initial capital letters. Other classes of words may also be considered taboo. Hyphenated words are an obvious choice for some, but who is to be the arbiter of the various whims of this practice? There will be those who balk at archaic, obsolete, or profane words, as well as alternative spellings of the same word in different countries (behaviour-behavior). The most valid additions to existing lists will be the innumerable words in common usage which are not listed in dictionaries. These include those words which are omitted to save space (obvious inflections, standard prefixes and suffixes), but also new words which have yet to be included.

Some might suggest that foreign words are to be disallowed. Who will be the judge of this class? Apparently, the best nine-square so far constructed is of wholly French words; none of the known English nine-squares has a full set of indisputable words. Would the quest for the first perfect ten-square be considered as fulfilled on finding a non-English creation? The German language, with its widespread use of compound words, would seem to be an obvious contender. And would a square of words from a number of different languages (a macaronic square) be accepted? This would certainly increase the wordstock!

Were a fully-acceptable ten-square finally to surface, it would likely be preceded by a number of progressive betterments of earlier attempts by Dmitri Borgmann, Jeff Grant, etc. These interim solutions are most likely to arise from the power of the computer, which is being harnessed to the task by a number of formists. The first of any new offerings may be expected to contain some fairly abstruse words taken from the latest additions to the word-lists; with the tremendous effort devoted to the ten-square over the last century, it is hardly likely that a square of well-known words has been lurking unfound all this time.

Adding more words to a database, especially words which may not form any part of a successful outcome, will greatly exacerbate search time. But who can foresee which words are worth adding, and which should be ignored?

It would surely be a great disappointment to achieve a near-perfect solution, i.e., one which was rejected owing to the inclusion of one or more doubtful words. Would it not be prudent to establish rules now for a valid ten-square, in order to avoid wasted time and effort, and to share all effort and information in a desire to produce a result before the end of time? (One can anagram the title to "Ten-Nutters Unique Share", a motto for this effort.) This is the approach often employed in medical and scientific research; results are pooled in the common interest. While the search for the ten-square cannot be considered as important, formists should accept that there are other logologists, less-addicted to this particular craving, who would nevertheless appreciate seeing a successful outcome of their monumental task.

The first step in a cooperative effort would be to establish a master set, as complete as possible, of acceptable words. This could be produced from a coalescence of the word-bases of the various formists, which could then be pruned of unacceptable words (and repetitions eliminated). The next step would be to examine the various computer programs which have been developed and, either by accepting the best or by producing the optimum from them, to use this as the computing standard. This would require agreement as to the algorithms affecting the direction of working through the square, and whether the words in the database should be listed in normal or reverse order.

The establishment of the computer program could be aided by standard test runs on the present programs. Each program could be timed against a run through a given test square, either a concocted one or one of the current submissions, such as the Borgmann square. Ideally, the same database should be employed, whether in normal or reversed format, but in any case the set of words to complete the test square must be in the database used. The use of differing databases would make it difficult to draw conclusions as to the merits of the various programs submitted to those tests; the inclusion of a single different word in the database could significantly change the timing. Each program would be run to simulate both top-down and bottom-up, in order to take full account of their differences.