WORDS, NON-WORDS, NONCE WORDS

A. ROSS ECKLER
Morristown, New Jersey

"When I use a word," Humpty Dumpty said, in rather a scornful tone, "it means just what I choose it to mean -- neither more nor less."

"The question is," said Alice "whether you can make words mean so many different things."

"The question is," said Humpty Dumpty, "which is to be master -- that's all."

Lewis Carroll
Alice Through the Looking Glass

What is a word? From Lewis Carroll onwards, this apparently simple question has bedeviled countless word buffs, whether they are participating in a game of Scrabble or writing an article on long transpositions for Word Ways. The purpose of this article is to examine various aspects of the problem, in the hope that each person facing it may decide where to draw his own line between words and non-words. This article does not attempt to impose a single line which all logologists must toe. If logologist A chooses to use a larger stockpile than logologist B, it is a simple matter for the latter to ignore those words he does not allow. Thus, some logologists will aver that DETARRATED is the longest palindrome in the English language; others will prefer MALAYALAM; some will insist that DEIFIED or ROTATOR are the champions -- and all are correct, as long as they properly define their different word stockpiles.

To help the reader decide what constitutes a word, this article suggests a ranking of words in decreasing order of admissibility. How can such a ranking be accomplished? A logical way to rank a word is by the number of English-speaking people who can recognize it in speech or writing, but this is obviously impossible to ascertain (even a sample of people would be costly to interview, considering that many words beloved by logologists are known to fewer than one person in ten thousand). Alternatively, one can rank a word by its number of occurrences in a selected sample of printed material. E. L. Thorndike and I. Lorge's Teacher's Word Book of 30,000 Words, published in 1945, is the most ambitious study of this type.
I am aware of; they examined more than twenty million words from a wide variety of sources. H. Kucera and W. N. Francis's *Computational Analysis of Present-Day English* is based on one million words from sources in print in 1961. Unfortunately, the majority of the words in Webster's Unabridged do not appear even once in these compilations -- and the words which do not appear are the ones for which a philosophy of ranking is most urgently needed. Furthermore, the written ranking will differ from the recognition ranking; vulgarities and obscenities will rank much higher in the latter than in the former.

A detailed, word-by-word ranking is an impossible dream, but a ranking based on classes of words may be within our grasp. Echoing the title, I propose the following classes: (1) words appearing in one (or more) standard English-language dictionaries, (2) non-dictionary words appearing in print in several different contexts, (3) words invented to fill a specific need and appearing but once in print.

Most people are willing to admit as words all uncapitalized, unlabeled entries in (say) Webster's New International Dictionary, Third Edition (1961). Intuitively, one recognizes that words become less admissible as they move in any, or all, of three directions: as they become more frequently capitalized, as they become the jargon of smaller groups (dialect, technical and scientific), and as they become archaic or obsolete. As Dmitri Borgmann has pointed out in *Beyond Language* (Scribner's, 1967), these classes have no definite boundaries. Is a word last used in 1499 significantly more obsolete than a word last used in 1501? Is a word known to 100,000 chemists more admissible than a word known to 90,000 Mexican-Americans? Each person must set up his own boundaries, recognizing their arbitrariness; to avoid endless hassles, it is convenient to set one's boundaries in terms of the labels one or another of the standard dictionaries uses. For example, one may admit words labeled rare but not obs. and words labeled sometimes cap. but not usu. cap. Or, one may allow all obsolete words in Webster's 2nd (excluding words last used before 1500) instead of Webster's 3rd (excluding words last used before 1755).

The second class consists of non-dictionary words appearing in print in a number of sources. There are many non-dictionary words in common use; some logologists would like to draw a wider circle to include these. Such words can broadly be classified into three groups: (1) neologisms and common words overlooked by dictionary-makers, (2) geographical place names, (3) given names and surnames. Let us consider these in turn.
In *Language on Vacation* (Scribner's, 1965), Dmitri Borgmann points out that the well-known words UNCASHED, EX-WIFE and DUTY-BOUND appear in no dictionaries (since then, the first of these has appeared in the Random House Unabridged). Few people would exclude these words from use in logological studies. Neologisms present a slightly more awkward problem, since some of them may be so ephemeral that they never appear in a dictionary. Perhaps one should heed Pope's dictum "Be not the first by whom the new are tried / Nor yet the last to lay the old aside."

Although geographical place names appear in the gazetteer sections of dictionaries (and the main section of the Random House Unabridged), much larger treasure-troves of place names can be found in *The Times Atlas of the World* (over 200,000 in either the five-volume Mid-Century Edition or the one-volume Comprehensive Edition), and the Rand McNally Commercial Atlas and Marketing Guide (over 100,000). However, these words are not all different, as in dictionaries; furthermore, some place names are already dictionary words. All these non-dictionary words can be easily verified by other readers; however, some logologists will feel uneasy about admitting as a word the name of (say) a small Albanian town which possibly has never appeared in any English-language text other than atlases. On the other hand, place names such as GUATEMALA are well-known and not lightly discarded.

Given names appear in the appendix of many dictionaries. Common given names, such as EDWARD and CORNELIA, ought to be admitted as readily as common geographical place names. However, the set of given names does not add much to the logological stockpile.

Family surnames at first blush appear to be on the same footing as geographical place names. However, one must be careful about sources. Biographical dictionaries, encyclopedias and Who's Who are perfectly adequate references for citing the existence of a surname, but one should be more cautious about citing surnames which appear only in telephone or city directories. Although (as Dmitri Borgmann has noted) these people are just as real as the ones appearing in the former references, once a telephone directory is supplanted by a later edition, it is very difficult to locate copies of it for verifying surname claims. Furthermore, telephone directories are not immune to nonce names, coined by subscribers for personal reasons (such as a desire to be the first or last listing). A good index of the relative admissibility of surnames is the number of people in the United States bearing that surname. An estimate of this could be obtained from the computer tapes of the Social Security Administration; in 1957, they issued a pamphlet giving the number of Social Security accounts associated with each of the 1,500 most com-
Clrgmann E and first of people. Perhaps he new

other sec-

house Un-

be found

e five-

ide (over

ary words.

iter read-

ning as a

has

es. On

known

ists. Com-

to be ad-

owever,

stockpile.

a footing

about

's Whos

 of a sur-

es which

Dmitri

mes ap-

ry is sup-

es of it

rectories

personal

A good

er of

ate of

l Security

number of

mon family names. At that time, the total number of distinct sur-

ames in their files was well over one million, a stockpile consider-

ably exceeding Webster's 2nd. If this list were ever published, it

would be of great logological interest.

The third and final class of words consists of nonce words --
those invented to fill a specific need, and appearing only once (or
perhaps only in the works of the author favoring the word). Few

logologists feel comfortable about admitting these as words.Nonce

words range from coinages by James Joyce and Edgar Allan Poe

(X-ing A Paragraph) to interjections in comic strips (Agghh! Yow-

ie!). Misspellings in print should probably be included also.

In the essay entitled "From Word to NonWord" in Beyond Lan-

guage, Dmitri Borgmann proposes that the logologist be prepared
to admit words that may never have appeared in print. For example,

he points out that Webster's 2nd lists EUDAEMONY as well as the

eye "Eudalmonia, eudalmonism, eudalmonist, etc. Variants of

Eudalmonia, eudalmonism, etc." From these he concludes that

EUDAEMONY must exist and should be admitted as a word. In sim-

ilar vein, he can conceive of sentences containing the word GRA-

CIOUSLY'S ("There are ten graciously's in Anna Karenina") and

SAN DIEGOS ("Consider the luster that the San Diegos of our nation

have brought to the United States"). In short, he argues that these

words might plausibly be used in an English-language sentence, but
does not assert that he knows of any actual usage. His criterion for

the acceptance or rejection of a word is its logological uniqueness;

he is far less likely to reject a word which is an eleven-letter palin-
drome than one which is a six-letter palindrome. (EUDAEMONY is

a relatively short word containing all five vowels and the letter Y.)

If words are to be admitted by logological uniqueness rather than

established usage, it is only a short step for a logologist to coin

words with any desired property -- say, a thirty-letter palindrome.

Doing one's own thing may be fashionable (or pop) psychology, and

may even lead to great art, but it is bad logology. Despite Humphry

Dumpty's assertion, words should still be our master -- an external

reality which the recreational linguist diligently examines, hoping to

discover unsuspected and captivating patterns.