THE LONGEST DICTIONARY WORDS

A. ROSS ECKLER
Morristown, New Jersey

In Language on Vacation (Scribner's, 1965), Dmitri Borgmann devotes several pages to a survey of the longest words in the English language. As might be expected, he finds it necessary to proceed from standard reference works such as the Webster dictionaries to ever-more-esoteric sources as the word-length is increased. One rapidly enters the realm of the nonce word -- a word used but once, or at most a few times by a single author. Since Borgmann's book appeared, it has become apparent to logologists that chemical names describing complex organic compounds can be almost infinitely expanded; the Guinness Book of Records cites a 1913-letter chemical term which earlier (May 1968) was published in Word Ways. Yet, it is again likely that one is entering the realm of the nonce word; in writing about a compound, a chemist is more likely to refer to it by formula if the name is at all long.

What are the longest words in the English language that are really used? One ought to be able to turn to unabridged dictionaries to answer this question; presumably they are interested in documenting language actually in use instead of flattering the vanity of word-coiners. Did not the compilers of Webster's Third proudly point out how many millions of examples from a multitude of sources they collected and collated? Is it not true that scores of such citations appear on every page? Alas, it soon becomes evident that such care has been lavished on only the short and medium-length words. I do not recall seeing a single citation illustrating the use of any word of twenty-three letters or more (the cutoff point for this article).

However, the most damning evidence that dictionary editors have relaxed their vigilance when considering really long words is contained in the Seventh Edition of Webster's Collegiate Dictionary: the longest word therein is our old miner's lung disease friend, PNEUMONOULTRAMICROSCOPICILICOVOLOXANOCOISIS, but the second longest word, ELECTROCARDIOGRAPHICALLY, is only half this length. The suspicion is inescapable that Webster's has inserted the former word in the Collegiate for reasons other than the everyday linguistic needs of the college student; can any reader cite a single legitimate source-use for this word? (I rule out the host of sources which discuss this as a long word, and instead seek ones that use it in context.) Funk & Wagnalls ruthless excises the middle of the word and reports only PNEUMOCOISIS; the same spelling is used on page 26 of Time Magazine, February 28, 1972.
Although it may cast little light on usage, a study of long dictionary words should illuminate other questions. How much consensus is there among different unabridged dictionaries concerning which words to include? How much change is there in the stock of long words over several decades? What classes of long words do dictionary-compilers favor?

The list at the end of this article is the result of a fairly thorough search through five different twentieth-century American unabridged dictionaries: Webster's New International Dictionary of the English Language (First Edition 1909, Second Edition 1934, Third Edition 1961), The Random House Dictionary of the English Language (1966), and Funk & Wagnalls New Standard Dictionary of the English Language (1946). The list is restricted to single-word, unhyphenated entries of 23 letters or more in boldface type. No plurals or other derived forms have been included unless specifically listed; however, in one case a single word from a two-word entry has been included. Variant spellings of the same word are separately listed. After each word, the source(s) in which it appears is given in code (l, 2, 3, R, F); asterisked sources refer to the Webster Addenda (published in 1929, 1957 and 1971) or the Funk & Wagnalls Addendum (published in 1956). Each word is classified by field: Ch for chemistry, Me for medicine, Sc for other scientific usage, and Li for literary usage.

Why an arbitrary cutoff point of 23 letters? It was selected partly to keep the list from growing too long, but more as a response to the claim in Funk & Wagnalls that the 22-letter word HONORIFICABILITUDINITY is "frequently cited as the longest word in the English language". (Even Funk & Wagnalls does not believe this, for it lists at least four longer words.)

What can be concluded from the list given below? Long words are overwhelmingly technical in nature; of the 77 listed, 29 are chemical compounds, 24 are medical terms, and 11 refer to other fields of science and technology. There are only 13 words -- less than twenty per cent -- that can be called literary, and of these several were originally coined as nonsense words. (Readers will recognize the most recent one, from the Walt Disney movie based on Mary Poppins; but others have a more distinguished paternity, such as Shakespeare and Gladstone.)

It is much more instructive to analyze the changes that have occurred in the Ch-Me-Sc-Li mix of long words over the past two generations. Consider the three editions of Webster's:

<table>
<thead>
<tr>
<th>Edition of Dictionary</th>
<th>Number of Words</th>
<th>Li</th>
<th>Sc</th>
<th>Me</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Edition (1909-29)</td>
<td>7 6</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Second Edition (1934-57)</td>
<td>28 5 1</td>
<td>4</td>
<td>6</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>Third Edition (1961-71)</td>
<td>29</td>
<td>1</td>
<td>3</td>
<td>11</td>
<td>14</td>
</tr>
</tbody>
</table>

The numbers spaced midway between the rows refer to words appearing in both the First and Second Editions (6), and both the Second and
Third Editions (5). Only one word, PSEUDOLAMELLIBRANCHIATA, appears in all three Webster dictionaries. The twentieth-century explosion of science and technology is mirrored in the ever-increasing total for the successive editions, but is shown even more clearly by the shift in emphasis from literary words in the First Edition to medical and chemical words in the Third.

Another way to look at the list is to see how much agreement there is among different dictionaries; do different editors apply similar criteria for including or excluding words? To minimize the secular increase demonstrated above, three reasonably contemporaneous dictionaries were selected: Funk & Wagnalls with Addendum (1956), Webster's Third without Addendum (1961), and Random House (1966). It is obvious that the Funk & Wagnalls editors are far more conservative about admitting long words than are either the Webster or the Random House editors; the numbers allowed are 9, 27 and 30, respectively. More startling, only 8 words (about 25 per cent of the total in either dictionary) are found in both Webster's and Random House, 2 in both Random House and Funk & Wagnalls, and none in both Webster's and Funk & Wagnalls. Such profound disagreement shakes one's confidence in the omniscience of dictionary-compilers; apparently each dictionary has its own preferred sources.

One rather obvious source for medical terms is Dorland's Medical Dictionary, brimful of sesquipedalian medical and chemical terms. In the 23rd Edition, there are 44 that are 25 or more letters long, including the 37-letter HEPATICOCOLONICCHOLOEYCESTENTEROSTOMY. However, of the 25 chemical or medical terms of 25 or more letters appearing on the dictionary list below, only six are found in Dorland.

What is the attitude of British dictionary-compilers toward long words? The Oxford English Dictionary, to the best of my knowledge, lists only two boldface unhyphenated words of 23 or more letters: ANTHROPOMORPHOLOGICALLY and TRANSUBSTANTIATIONALIST. (However, if hyphenated words are allowed, the OED also recognizes the 29-letter nonsense word FLOCCI-NAUCI-NIHILI-PILLIFICATION.) Although the original twelve volumes were compiled prior to the beginning of the century (hence, before the proliferation of scientific terminology), the Supplement was published in 1933, only one year before Webster's Second Edition. Apparently the OED subscribes much more rigorously than any of the American dictionaries to the philosophy that (literary) usage is the principal criterion for inclusion.

It is quite evident that unhyphenated words of 23 letters or more are rarely encountered by the average reader. In Computational Analysis of Present-Day American English (Brown University Press, 1967), Henry Kucera and W. Nelson Francis report on a sample of one million words taken from a varied collection of writings by American authors that were first published in the United States in 1961. Only four words of 20 or more letters turned up in this sample: INSTITUTIONALIZATION, PSYCHOPHARMACOLOGICAL, IMMUNOELECTROPHORESIS, and ALKYLBENZENESULFONATES. In their sample, 17-letter words occur about half as often as 15-letter ones and about half as often as 13-letter ones; the law-of-thumbs for 20-letter words is about half as often as 15-letter ones and fewer hyphenated words.

Readers of these dictionaries alter the basic behavior of readers because they are used to proper names.
RANCHIATA, a century earlier increasing clearly by
seemment thereby similar cri-
secular in-
1(1956), House (1966). re conserva-
 or the Ran-
le total in
h House, 2 both Web-
eters; appar-
and's Medi-
ers long, in-
ENTER-
 of 25 or
six are found
ward long
knowledge,
letters:
ATIONALIST.
o recognizes
FICATION.)
to the begin-
entific term-
year before
much more
s or more
ional Analysis-
Press, 1967). of one million
ican authors
ly four words
ATIONALIZA-
PHORESIS
-letter words
occur about half as often as 16-letter words, 18-letter words occur
about half as often as 17-letter words, and so on; extrapolating this
rule-of-thumb, one can expect about one 25-letter word in a sample
of 15 million words, or one 30-letter word in a sample of half a bil-
ion words (ten sets of the Oxford English Dictionary). However,
Kucera and Francis's index heading on page 129 succinctly com-
ments on such extreme extrapolations: unutterably - wacky.

Readers are invited to discover words that I have overlooked in
these dictionaries; however, it is unlikely that such discoveries will
alter the basic thrust of this article. As a more ambitious project,
readers are encouraged to extend the study to hyphenated words, or
to proper names (using the Times Atlas and the Rand McNally Com-
mercial Atlas, for example). To start the ball rolling, here are a
few hyphenated words that were noted in passing:

- quasi-internationalistic
- quasi-extraterritorially
- squarrose-pinnatipartite
- ecclesiastic-co-conservative
- chrysanthemum-dicarboxylic
- vajrachchedika-prajnaparamita
- interstitial-cell-stimulating
- hystero-salpingo-oophorectomy
- hyperglycemic-glycogenolytic

I am indebted to Darryl Francis for communicating to me the con-
tents of the Air Force reverse word list (all words of 25 letters or
more are listed separately). Both he and Ralph Beaman suggested
various long words to be added to the list.
45 pneumonoultramicroscopicsilicovolcanoconiosis
pneumonoultramicroscopicsilicovolcanokoniosis
3 R Me
2*R Me

34 diaminodipropyltetramethylenediamine
supercalifragilisticexpialidocious
R Ch
R Li

31 dichlorodiphenyltrichloroethane
electrocardiographically
R Ch

29 cyclotrimethyleneetrinitramine
trinitrophenylmethylnitramine
R Ch
3 R Ch

28 antidistrust establishmentarianism

27 electroencephalographically
ethylenediaminetetraacetate
honorificabilitudinitatibus
hydroxydeoxytocicosterone
octamethylpyrophosphoramide
3 R Me
3 2 Me
1 2 Li
3 Ch
3 Ch

26 anhydrohydroxyprogesterone
cystoureteropyelonephritis
ethylenediaminetetraacetic
ethylmercurithio salicylate*
hydroxydeoxytocicosterone
pancreaticolecystostomy
phenoltrichlorophthalein
pneumoencephalographically
3 Ch
3 Me
3 Ch
3 Ch
2 Me
2 Ch
2 Me
3 Me

25 demethylchlortetracycline
electrocardiographically
electroencephalographical
formaldehyde sulphoxylate
immunoelectrophoretically
occipitoresorciptoscapular
philosophicpsychological
superincomprehensibleness
syngenesistransplantation
tetra bromophenolphthalein
3 R Me
1*R Me
1*2 Me
3* Me
F Me
2 Sc
R Li
3 Me
2 3 Ch

24 Aldiborontiphoscoporphine
dioxydialmidobenzenobenzol
diphenylaminechlorarsine
dis establishmentarianism
electrocardiographically
haematospectrophotometer
laryngotracheobronchitis
magnetothermoelectricity
meningoencephalomyelitis
pathologicpsychological
pseudointernationalistic
1 F Li
F Ch
3 R Ch
3 R Li
3 R Me
2 Me
3 Me
R Sc
3 Me
2 Sc
R Li

23 anthroblepharon
dichlorodimethylelectroformalgastroheteromeshydrhonificabilitudinitatibus

desoxy
dichloro
dimeth

electro

electro

formal
formal

overin
overin

parasy
parasy
pathol
pathol

perica
perica

phenol
phenol

philo
philo

polytet
polytet

Pseudo
Pseudo

pseudo
pseudo

scient
scient

thymol
thymol

transu
transu

**prec
**prec
scientifically philosophical
tetraiodophenolphthalein
thyroparathyroidectomy

23 anthropomorphologically
blepharosphincterectomy
chlorotrifluoroethylene
desoxyribonucleoprotein
dichlorodifluoromethane
dimethylthiocarbamate**
electroencephalographic
electrophotomicrography
formaldehyde sulfoxylate
formaldehydesulphoxide
gastroenteroanastomosis
heterocomplementophilic
magnetohydrodynamically
methylidihyromorphinone
microspectrophotometric
occipitotemporoparietal
overindividualistically
overintellectualization
pancreaticoduodenostomy
parasympatheticomimetic
pathologicohistological
pericardiomediastinitis
phenolsulphophthalein
phenoxybenzylpenicillin
philosophicotheological
polytetrafluoroethylene
Pseudolamellibranchiate
pseudolamellibranchiate
pseudophilanthropically
scientificcogeographicall
thymolsulphonephthalein
transsubstantiationalist

*preceded by the word sodium
**preceded by the word ferric