

# Current Comments

## Another Look at ISI's Master Dictionary— Aiding Scientific Etymology and Reflecting Changes in Science

Number 43

October 26, 1981

*Anarchy* increased between 1976 and 1980—but so did *ethics* and *serendipity*. *Feminists* almost doubled during the same period, and *sexism* declined dramatically. As *beef* dipped, *breastfeeding* and *cannibalism* increased. Both *constipation* and *diarrhea* increased, as did *cancer* and *herpes*. But *anti-cancer* increased tenfold, and *tumors* fell off. *Promiscuity* and *erotica* both increased, but *lust* faded and *sexuality* sagged. Things were looking up for *evangelists* and *androgyny*. *Violence* and *terrorism* both increased, but so did *happiness* and *optimism*, while *backaches* and *fatheads* became less numerous.

Has anarchy really increased? Have instances of breastfeeding and cannibalism actually risen? Only in the sense that from 1976 to 1980, occurrences of these words in the titles of articles indexed by ISI® increased. Increases in the occurrence of certain words, such as *song* and *Tolkien*, can be attributed at least in part to the addition of the *Arts & Humanities Citation Index™* to our product line.<sup>1</sup> But it cannot be denied that the vocabulary of science, like that of any living language, is constantly changing. New words are coined to describe new substances, improved processes, or previously undiscovered phenomena, while old words fall into disuse. Changes in the activities of scientists are thus reflected in the words they use—and in the words they choose for the titles of their articles.

In an essay published several years ago, we examined the apparent shifts in

scientific interest that occurred between 1973 and 1976.<sup>2</sup> We were able to do this because ISI uses title words, as well as other bibliographic descriptors, to index each article added to our data base. We can therefore systematically observe and quantify changes in the activities of scientists, based on the frequency of occurrence of the words they use.<sup>3</sup> How accurately this practice reflects scientific activity has never been qualitatively determined.

What makes the monitoring of the ebb and flow within the language of science possible is the Unique Word Dictionary (UWD). The UWD is one of several manual and computer quality-control routines we use to verify each title word before it becomes part of our data base.<sup>4,5</sup> The UWD is a machine-readable master list of correctly spelled terms that have been certified as "real" words. Real words are those that have been verified as accurate and authentic by our editors. It was the information in the earlier essay, and the light it threw on where science is taking us, that was featured in an article by Dan Greenberg in the *Washington Post*.<sup>6</sup>

Table 1 contains 196 terms and some of their spelling variants, selected from the UWD. To the right of each word is listed its frequency of occurrence in article titles for the years 1973, 1976, 1979, and 1980. Many of these words appeared on a similar list in the previous essay. If you look at that essay, you may note the frequencies given in it do not match the figures given in this update.

For example, *backache* was listed as having occurred six times in 1973 and 14 times in 1976 in the previous essay. But Table 1 of this essay shows that *backache* occurred with a frequency of five and 14 times in 1973 and 1976, respectively. The reason for the difference is that we have changed our procedures for compiling the UWD since 1976. The new methods give more accurate counts.

Whenever a new article is added to our data base, the words appearing in

the title are automatically compared with the words already contained in the UWD. If any new word does not match a word already in the UWD, it is "kicked out" for evaluation by our editors. Many nonmatching words turn out to be simple keying or spelling errors, while a few, such as *aganglionosis*, are extant words that can be found in standard reference works, but for one reason or another have never before found their way into the UWD.

**Table 1:** Selected words from ISI's Unique Word Dictionary (UWD), which includes all words used in the titles of articles covered by ISI's data base. Each word's frequency of occurrence for 1973, 1976, 1979, and 1980 is indicated. The spelling variants listed in parentheses behind each word occur separately in the UWD, and are accompanied by their own unique frequency counts. For the purposes of this essay, however, the frequency counts of these spelling variants have been combined with the count for the main root. British spellings have been Americanized.

WORD	1973	1976	1979	1980					
adnamycin	209	1,020	1,845	1,763	constipation	47	34	103	75
aerosol (s) (ized)	1,922	2,355	2,587	2,300	copyright	51	198	275	137
aganglionosis	2	3	3		cosmogenic	15	13	56	43
algorithm (ic) (s)	2,060	2,762	3,221	3,574	counter-immunoelectrophoresis	2	4	6	8
alpha-fetoprotein	358	795	796	787	counterimmunoelectrophoresis	49	141	212	171
amniocentesis	85	117	288	263	cryo-ultramicrotomy	11	18	13	25
anarchy (ist) (ists) (ism)	44	96	152	168	cryoultramicrotomy	1	3	6	6
androgyny (ous)	1	40	158	100	cybernetic (s)	108	212	133	105
angiotensin	667	951	1,078	1,063	Decembrist (s)	1	35	22	10
angiotensin-I	27	16	44	36	deoxyribonucleic (DNA)	12,224	11,349	14,517	15,713
angiotensin-II	86	121	243	217	diabetic (s) (es)	4,068	5,637	7,069	9,235
anti-androgen (s)	38	29	111	185	diarrhea	672	868	1,068	1,412
anti-cancer	21	56	581	561	dienone (s)	42	14	56	46
anti-Parkinsonian	18	9	54	20	diiodoacetyl	0	0	6	86
anti-platelet	32	62	250	195	disappointment (ed) (ing)	165	254	238	401
antiandrogen (s)	28	19	0	0	divorce (s) (ces) (ed)	10	34	74	86
anticancer	29	38	0	0	DNA-binding	63	71	53	51
antiParkinsonian	6	7	0	0	DNA-dependent	396	196	261	246
antiplatelet	5	10	0	0	dopa	1,503	3,274	4,871	5,178
apricot (s)	47	51	134	38	dopamine	51	77	67	59
apricot-tree	0	0	1	3	dopamine-beta-hydroxylase	0	42	554	480
Armageddon	0	8	9	1	endorphin (s)	0	190	781	1,108
auto-antibody (ies)	0	12	13	35	enkephalin (s)	63	156	127	165
autoantibody (ies)	63	85	129	132	erotic (a) (ism)	595	1,426	1,636	2,109
backache (s)	5	14	9	7	euthanasia	23	106	93	83
beef	1,586	1,405	1,380	1,188	evangelism (ist) (ists)	9	8	46	22
beef-heart	1	15	48	32	exercise (s) (ing)	3,671	4,622	7,212	7,912
beta-lipotropin	3	65	103	191	fathead (s)	73	92	75	71
biofeedback	109	298	489	534	feminism (ist) (ists)	62	231	421	420
biomass	165	432	1,244	1,106	fiber-optic (s)	44	93	94	117
bionic (s)	0	7	10	4	fission	2,227	1,560	1,704	1,595
biorthym (s) (ic)	19	40	84	79	fundamentalism (ist) (s)	3	17	27	78
bombesin	23	102	197	298	fusion	1,729	2,789	3,054	3,420
breast-feeding	10	26	43	46	glycoprotein (s)	1,772	2,061	3,848	4,183
breastfeeding	11	16	25	37	gonadotropin	1,814	1,342	1,182	896
calmodulin	0	0	211	1,003	Gondwanaland	21	10	23	36
calmodulin-dependent	0	0	2	10	hadron (s)	583	750	734	624
cancer	3,863	6,476	7,059	6,956	handicapped	398	664	809	886
cannabis	53	50	35	42	happiness	29	105	145	120
cannabis-sativa	321	299	219	249	herbicide (al) (s)	3	1,060	2,268	1,623
cannibal (s) (ism) (istic)	41	94	117	120	herpes	312	403	513	623
cardiomyopathy	374	513	1,169	1,095	herpes-simplex	142	301	354	395
cephalosporns	158	233	366	350	herpes-virus	21	18	32	19
charm (ed)	25	546	501	373	herpesvirus	105	89	84	92
chocolate	43	26	137	153	histamine	42	2,370	2,855	2,960
cholecystokinin	241	306	481	544	homosexual (s) (ality)	216	323	355	428
chorio-gonadotropin	0	106	0	8	horseradish	0	76	119	149
choriogonadotropin	0	40	17	0	hybridomas	0	0	161	231
cimetidine	0	51	343	373	immuno-cytochemical	0	1	0	162
cimetidine-induced	0	0	7	11	immunocytochemical	0	14	8	44
citation (s)	26	41	328	339	immunomicrospheres	0	0	4	11
clone (s) (ing) (ed) (al)	1,222	1,701	4,056	5,825	inhibit	0	13	61	132
cocaine	293	419	438	335	insemination	122	17	216	249
cocaine-induced	0	1	6	2	interferon	1,187	1,300	2,163	2,661
					laetile	1	34	80	51

L-dopa	1,160	841	661	383	rheumatic (toad)	1,026	1,028	1,358	1,260
lemming (s)	21	65	96	73	ribonucleic (RNA)	6,165	5,858	5,579	5,409
leprosy	340	510	1,535	1,017	ringspot	187	73	123	86
lepton (s)	359	445	288	456	rural	1,338	2,145	2,792	3,009
linguistic (s)	625	1,581	1,500	1,737	rural-urban	1	3	3	2
lust	0	15	15	11	saccharin (ei)	165	134	275	298
macrophage (s)	2,204	3,591	5,444	6,429	sarcoplasmic	572	56	102	126
macrophage-lymphocyte	1	4	27	8	science	3,019	4,338	4,927	4,411
maize	286	305	432	607	scientific	1,649	2,953	3,334	3,471
marijuana (marihuana)	614	544	412	278	scintigraphic (y)	586	1,410	1,210	2,096
mastitis	223	257	547	482	sclerosis	277	456	632	682
melanocyte-stimulating-hormone	123	91	95	76	scurvy	19	47	33	18
melanin	329	434	758	708	secretin	530	863	610	731
metallo-porphyrin (s)	0	0	2	1	self-help	46	173	120	186
metalloporphyrin (s)	63	130	148	180	self-image	15	95	79	52
methylation	656	705	1,197	1,166	semantic (s)	624	1,136	1,106	1,135
micro-circulation	0	3	3	81	serendipity (ous)	8	11	46	39
micro-surgical	2	1	1	2	sexism	16	104	90	8
micro-surgery	1	0	75	56	sex-role	81	288	639	569
microcirculation	197	440	663	52	sexual	1,866	2,523	3,182	3,101
microsurgical	69	193	375	306	sexuality	155	432	324	383
microsurgery	19	29	0	0	sickle	22	37	44	37
mitogen (ic) (s)	680	1,279	1,951	1,916	sickle-cell	631	743	829	826
monocyte (s)	197	734	1,366	1,632	simplex	155	144	231	218
monopole	100	224	241	205	smog	101	91	113	109
morphometric	364	542	939	986	socialism (istic) (ist) (ists)	748	2,241	2,242	1,925
murine	2,624	4,604	7,575	8,301	sociobiology	0	53	84	107
murine-leukemia	0	0	0	2	soliton (s)	80	407	609	747
myelo-blastic	0	0	0	4	somatomedin	85	310	566	389
myeloblastic	51	206	53	18	somatomedin-A	0	3	16	1
myeloid	398	895	536	607	somatomedin-C	0	2	13	0
Neisseria gonorrhoeae	159	341	550	683	somatostatin	38	1,145	1,631	2,087
neurochemical	98	273	396	384	song (s)	195	1,354	1,292	1,410
neuropeptide (s)	0	0	183	264	spectrin	0	89	230	266
neurotoxicity	61	100	293	447	squamous	147	267	298	562
noradrenergic	238	468	667	770	squamous-cell	57	83	132	151
nude	178	845	353	258	stereopsis	16	72	66	57
nude-mice	0	0	137	140	steroidogenesis	327	256	599	529
nude-mouse	0	0	17	24	sub-population (s)	0	4	263	282
occult (ism)	196	310	452	559	subpopulation (st)	213	953	0	0
olefin	385	586	553	502	sulfobromophthalein	96	206	83	63
opiod (s)	19	47	672	798	superconductor (s) (ivity)	1,745	1,457	1,941	1,677
optimism (ist) (ists) (istic)	45	94	99	124	T-cell (s)	412	1,824	5,638	8,694
paracetamol	123	293	538	414	Tagamet	0	0	1	1
parity	365	501	543	639	terrorist (ism) (st)	26	111	306	262
parity-violating	6	6	14	16	thalidomide	94	64	67	98
pharmacokinetic (s)	291	711	1,085	1,204	tokamak (s)	543	1,339	1,527	1,785
phosphorylase	683	789	697	470	Tolkien	4	19	13	58
phosphorylase-A	4	5	12	9	transsexual (s) (ism)	24	76	60	61
phosphorylase-B	13	27	31	15	trivia (ial) (iality)	38	147	132	60
plasmid (s)	381	1,150	2,936	3,286	tumor (s)	10,592	14,148	13,061	13,991
positron (s)	361	634	1,051	1,124	tumor-associated	30	58	71	75
prolactin	2,283	3,530	5,185	5,194	tumor-bearing	41	62	91	100
promiscuous (ity)	11	33	10	16	tumor-stimulating	0	0	1	0
prostaglandin (s)	2,204	6,477	6,160	6,804	ultrasonography	62	244	657	786
protein kinase	0	852	1,741	1,892	ultrasound	1,005	1,241	2,182	2,611
pseudohypoparathyroidism	23	59	60	116	vanadate	91	77	509	475
Purkinje	21	17	13	8	vasopressin	712	1,163	2,098	1,897
Purkinje	252	341	505	585	Viet-Nam	3	0	0	0
quark (s)	269	570	1,261	1,233	Vietnam (ese)	401	400	497	826
quark-antiquark	1	1	10	11	Vietnam-War	14	36	31	41
radioimmunoassay	1,899	3,042	3,423	3,717	violence (violent)	388	741	987	858
recombinant	16	49	131	132	vitellogenin	20	169	264	280
renal	7,995	10,305	12,221	11,294	Voyager (s)	0	1	95	76
renin	953	1,931	2,313	2,021	Voyager-1	0	0	25	8
renin-aldosterone	3	8	17	14	Voyager-2	0	0	16	5
renin-angiotensin	181	384	483	466	winter	840	1,494	2,175	2,186
renin-angiotensin-aldosterone	6	3	40	4	woman (s) (en) (en's)	4,024	8,130	9,427	10,602
renomedullary	27	104	83	41	yogurt	27	63	96	94
reticulum	603	98	159	66	Zambia	122	249	222	184

Many words, however, can be classified as valid only after our editors have carefully reviewed the article and the context in which the word in question appeared. If a word seems valid based on this evaluation, and yet cannot be found in any reference work and has never before appeared in the UWD, it can cautiously be labeled new. Eventu-

ally, every word that has been tagged for further evaluation is either corrected and eliminated, or verified and added to the UWD.

Actually, the very name of the UWD illustrates one of the processes by which new words are sometimes coined. No one here at ISI can remember who first used the phrase "Unique Word Diction-

ary" to describe this particular computer-generated word list. Yet today it is an unquestioned part of ISI's in-house jargon. Apparently someone who worked with the UWD when it was first generated coined the name, others began to use it, and eventually UWD joined the rest of the acronyms in ISI's alphabet soup.

In 1980, approximately 450,000 different terms were used in the titles of articles we indexed. During 1979, the total was 430,000. These figures compare with the 1977 total of about 300,000 different terms. However, most of the increase over these years is due to spelling variations and extant words that, as mentioned earlier, never before appeared in the UWD. We estimate that only 500 to 1,000 words are newly coined and used for the first time each year.

Some new words are simply the result of new or unusual combinations of prefixes or suffixes with known roots, and their meanings are usually self-evident, as in the word *anti-Parkinsonian*. Other words are completely original and more often than not, their precise definitions are elusive without a proper explanation. For example, *calmodulin* is formed by combining the *cal-* from calcium with the *modul-* from modulator, plus the word *in*, which is used to signify a neutral chemical compound. But the exact meaning of *calmodulin*—a protein that acts as a receptor and intermediary for calcium ions in the regulation of numerous cellular activities<sup>7</sup>—would probably escape the average reader.

Another word in this category is *hybridomas*, "artificially-created cells that produce pure or 'monoclonal' antibodies."<sup>8</sup> The word did not occur at all in either 1973 or 1976, but appeared 161 times in article titles in 1979, and 231 times in 1980. As Nicholas Wade explains, hybridomas produce "a constant and uniform source of pure antibody, instead of the usual mixture produced

by the immune system, [affording scientists] a powerful research tool [that] can be expected to provide quicker and more accurate diagnoses of viruses, bacteria, and cancer cells."<sup>8</sup>

Other words that qualify as new consist of combinations of familiar terms that only make sense after you've been told what they mean. *Immunomicrospheres*, for instance, are "specially designed microscopic particles that have antibodies or similar molecules chemically bound to their surfaces. The antibody-coated microspheres react in a highly specific way with target cells, viruses, or other antigenic agents."<sup>9</sup> On the other hand, *cosmogenic*, which is in the same category (although it certainly cannot be called a *new word*), requires no such elaborate explanation. It refers to effects produced by cosmic rays.

Apart from words that occur for the first time, such as *choriogonadotropin*, *dioctadecyl*, *endorphin*, and *enkephalin*, the UWD also includes already existing words that are being used in new ways, whether alone or in combination with other words. Of course, new uses for old words probably cause more grief for ISI's indexing systems than the use of idiomatic expressions. Actually, idiomatic expressions are often less ambiguous than the use of an old word in a new way because the latter can be identified only by examining the context in which the word appeared. For instance, deciphering the idiomatic expression *on the beam* is far less troublesome than deciding whether the word *beam* refers to a ray of light or one of the supporting struts of a roof.

It should be kept in mind when examining Table 1 that the variation in the frequency of certain words may well be the result of changes in ISI's product line, as well as the addition or deletion of journal and non-journal material to our data base. For example, non-journal material, such as proceedings and books, was added to the *Science Citation Index*<sup>®</sup> (*SCI*<sup>®</sup>) in 1977,<sup>10</sup> and to the *Social*

*Sciences Citation Index*<sup>®</sup> (*SSCI*<sup>®</sup>) in 1979.<sup>11</sup> And since the previous essay, we started publishing *Current Contents*<sup>®</sup>/*Arts & Humanities*,<sup>12,13</sup> as well as the *Index to Scientific & Technical Proceedings*<sup>®</sup>,<sup>14</sup> and the *Index to Social Sciences & Humanities Proceedings*<sup>®</sup>.<sup>15</sup> Incidentally, the high counts recorded in 1979 and 1980 for the term *citation* is self-generated, since the UWD of course includes *Current Contents*.

Between 1973 and 1976, the number of source articles in the ISI data base increased by approximately 30 percent. Between 1976 and 1979, the increase was almost 50 percent, from 615,000 articles to 895,000. But in 1980, the increase was just 100,000 articles, or about ten percent. Since the size of the data base grows every year, affecting the frequency of many words, we chose to include in our survey mainly those words whose frequencies changed by 50 percent or more over the period between 1976 and 1980, though we have also included other words of special interest, whose frequencies did not change much, if at all.

It is difficult, however, to assess the impact of data base changes for every word because of uncertainty surrounding the use of particular words. Such a word is *horseradish*, which jumped from no occurrences in 1973 to 76 in 1976, 119 in 1979, and 149 in 1980. We strongly suspect that this rise reflects an increase in the use of the compound horseradish peroxidase, which is a kind of stain used in microscopic research. But without actual examination of all the literature in which the word *horseradish* appears, we cannot be certain that it—or any other word, for that matter—is always being used in a way we would expect. *Horseradish* might just as easily refer to the horseradish flea beetle, *Phyllotreta armoraciae*, or even to the familiar condiment that you might enjoy with a sandwich.

The case is much the same for words like *charm* and *nude*. Occurrences of

the word *charm* in article titles rose from 25 in 1973 to 546 in 1976, then dropped to 501 in 1979 and to 373 in 1980. *Nude* occurred 178 times in 1973, 845 times in 1976, 353 in 1979, and 258 in 1980. The precise ways in which these words were being used, however, is not necessarily clear in all cases. *Charm*, for instance, was probably being used to refer to a certain property of subatomic particles. But conceivably, it could have concerned enchantments or lucky amulets. Similarly, *nude* was most likely being used in reference to the nude mouse, a hairless mutant that has become popular in cancer and immunology research. But obviously, *nude* might also have been used in talking about sun worshipers, or about the titles of paintings or photographs.

Incidentally, the word *nude* allows me to point out another interesting facet of the UWD. If you check the table, you will note that below *nude*, the hyphenated terms *nude-mouse* and *nude-mice* are listed. Whenever a term in the UWD begins to be paired frequently and consistently with another term, our editors program the computers to record a frequency count not only for those words when they occur separately, but also when they occur together as a pair. In this case, *nude* was being paired often enough with *mouse* and *mice* to warrant just this kind of action. The UWD continued to record occurrences of *nude* and *mouse* and *mice*, of course, when those words occurred separately or in combination with other words. But when they occurred together with one another, the count was recorded under the "word pair," *nude-mouse*, or under the word pair, *nude-mice*.

Obviously, this sort of treatment can have an effect on the total frequency count of a word. The totals under the years 1979 and 1980 for *nude* show all the occurrences of *nude* in article titles for those years—*except* when it appeared together with *mouse* or *mice*.

Those occurrences are recorded under the special entries *nude-mouse* and *nude-mice*. To determine the true total of the number of times *nude* appeared in article titles for those two years, you would have to add the totals for all the word pairs the word appeared in to the totals for the word itself.

An especially dramatic example of what can happen when the way a word is keyed into the UWD is changed is that of *subpopulations*. This word in 1973 and 1976 was entered without a hyphen. But between 1976 and 1979, the spelling was changed to *sub-populations* in order to reflect changes in the way the word was being used by scholars, and to make it easier to conduct a search for the word. The counts for the old spelling of the word—which was retained in the UWD—dropped to zero for both 1979 and 1980, while the counts for the new spelling were 263 and 292, respectively.

So far I have talked about words that present obvious ambiguities for someone casually thumbing through a UWD printout. Some words in the UWD, however, are intriguing precisely because they are *not* ambiguous—or at least they don't seem to be at first glance. For instance, there aren't too many choices of connotation for a word like *fathead*, which occurred 73 times in 1973, 92 times in 1976, 75 times in 1979, and 71 times in 1980. That's a lot of appearances in prestigious scholarly journals for a word whose primary meaning is insulting. Are scientists taking potshots at one another in the titles of their articles? An examination of a number of the article titles in which *fathead* appeared, however, quickly dispels any such notion: in the majority of cases, it was being used to refer to the fathead minnow.

Similarly, *Armageddon* may seem an odd subject for a scholarly article. But a quick search of the literature reveals that it wasn't appearing in the titles of

scholarly articles at all. Instead, it was being used in the headlines of letters to the editors of various journals. The omnivorous UWD makes no distinction between these titles and the titles of articles. And since ISI expanded its non-journal coverage to include books and book reviews (whose titles are generally more fanciful than the titles of scholarly articles), as well as journals in the arts and humanities, there are plenty of article titles in the UWD that aren't exactly "scientific."

The word *aganglionosis* allows me to point out something else about the UWD. *Aganglionosis*, according to the table, occurred only twice in 1973, three times in both 1976 and 1979, and seven times in 1980. These figures may not be entirely accurate, however. At the end of each calendar year, words which occur less than three times are automatically purged from the UWD, so that counts as low as those for *aganglionosis* are unreliable. This also means that words such as *endorphin*, *enkephalin*, and *neuropeptide*, which all registered zero occurrences in at least one column, may actually have appeared once or twice during that year, but failed to attain the minimum threshold level, and so went unrecorded. Thus, without actually consulting the *Permuterm*<sup>®</sup> *Subject Index*, we can't rely on the UWD to be precise about the year of a word's first appearance. But even if *neuropeptide*, for example, did appear a couple of times in 1976, the table shows that its use had increased dramatically by 1979.

Examining the list by broad subject headings, we see that in physics, *tokamaks*—a proposed design to contain and control the fusion reaction—increased from 543 in 1973 to 1,339 in 1976. It continued its upward climb in 1979 and 1980, reaching marks of 1,527 and 1,785, respectively. *Positron* almost doubled between 1973 and 1976, rising from 361 to 634, and almost doubled once again between 1976 and 1979, as it

ballooned upward to 1,051. Its rate of growth slowed by 1980, however, as it attained a mark of only 1,124.

In medicine, occurrences of *amniocentesis*, a surgical procedure for the detection of chromosomal abnormalities in fetuses, rose from 85 in 1973 to 117 in 1976. Its frequency more than doubled by 1979, with a total of 288, but dropped off to 263 in 1980. *Herpes*<sup>16,17</sup> recorded a steady increase, rising from 312 in 1973 to 403 in 1976, 513 in 1979, and 623 in 1980. The totals for *simplex* are erratic, however: it dropped from 155 in 1973 to 144 in 1976, jumped to 231 in 1979, then fell again in 1980 to 218. Meanwhile, the word pair *herpes-simplex* occurred 142 times in 1973, 301 times in 1976, 354 times in 1979, and 395 times in 1980. *Leprosy*,<sup>18</sup> which occurred 340 times in 1973, rose to 510 in 1976, and then leaped to 1,535 in 1979, though it dropped to 1,017 in 1980.

In immunology and biochemistry, *somatostatin*, the name of a chemical that inhibits the release of growth hormone, appeared in article titles only 38 times in 1973; by 1976, that number had inflated to 1,145 and continued to grow (though not so precipitously) in 1979 and 1980, when it attained marks of 1,631 and 2,087, respectively. *Dopamine* also registered impressive gains, though of a steadier variety: in 1973 it occurred 1,503 times, rising to 3,274 in 1976, 4,871 in 1979, and 5,178 in 1980. *Plasmid* rose from 381 occurrences in 1973 to 1,150 in 1976, 2,936 in 1979, and 3,286 in 1980.

Looking through some of the drugs on the list, it is interesting to note that *laetrile* rose from one occurrence in 1973 to 34 in 1976 and 80 in 1979, then fell to 51 in 1980. *Cocaine* occurred 293 times in 1973, 419 times in 1976, 438 times in 1979, and 335 times in 1980. *Interferon* reached marks of 1,187 in 1973, 1,300 in 1976, 2,163 in 1979, and 2,661 in 1980. *Cimetidine*, a new drug used in the treatment of ulcers, which I dis-

cussed in a previous essay,<sup>19</sup> climbed from no occurrences in 1973 to 51 in 1976, 343 in 1979, and 373 in 1980; meanwhile, *Tagamet*, a brand name under which cimetidine is marketed, registered frequency totals of zero in 1973 and 1976 and one in both 1979 and 1980.

In genetics, *clone* rose from 1,222 in 1973 to 1,701 in 1976, then more than doubled in frequency by 1979, registering a total of 4,056. By 1980, the figure had risen to 5,825. *DNA* went from 12,224 in 1973 to 11,349 in 1976, 14,517 in 1979, and 15,713 in 1980. Meanwhile, *RNA* totaled 6,165 for 1973, 5,858 for 1976, 5,579 for 1979, and 5,409 for 1980.

Among words that seem relevant mainly to the social sciences, the frequency of *anarchy* increased from 44 occurrences in 1973 to 96 in 1976, 152 in 1979, and 168 in 1980. *Divorce* increased from 165 in 1973 to 254 in 1976, dropped in 1979 to 238, then rose again in 1980 to 401. *Self-help* increased from 46 to 173 between 1973 and 1976, slid to 120 in 1979, then increased once again in 1980 to 186. *Sex-role* increased from 81 to 288 between 1973 and 1976, continued the upward trend to 639 in 1979, only to slip to 569 in 1980. *Woman* rose steadily throughout the study period, from 4,024 in 1973 to 8,130 in 1976, 9,427 in 1979, and 10,602 in 1980. *Feminism* increased from 62 occurrences in 1973 to 231 in 1976 and 421 in 1979, but hit only 420 in 1980.

Although changes in the number of times a word is used can give an indication of changes in scientific activity, more often than not, citation analysis—as I have repeated *ad infinitum*—may prove an even better indicator. We are currently performing experiments on the use of co-occurrence of words for just this purpose. Scholars have long studied the evolution of words and language as a window on the development of culture and society. In the same way, scientific etymology can

help us gain a better insight into the development of science. A dictionary of new scientific words, assembled with the help of the UWD, might prove an invaluable tool for researchers as well as publishers, librarians, students, and laypeople. Such a dictionary might be updated monthly to keep up with the changing vocabulary of science. All of this activity at ISI is coordinated by Jim

Shea and others who are working on the forthcoming *ISI Atlas of Science*, described in last week's essay.<sup>20</sup>

\* \* \* \* \*

*My thanks to Stephen A. Bonaduce and Edward M. Sweeney for their help in the preparation of this essay.*

©1981 ISI

## REFERENCES

1. **Garfield E.** Will ISI's *Arts & Humanities Citation Index* revolutionize scholarship? *Current Contents* (32):5-9, 8 August 1977.\*
2. ...., ISI's master dictionary aids scientific etymology and reflects changes in science. *Current Contents* (4):5-11, 23 January 1978.\*
3. ...., Jabberwocky, the Humpty-Dumpty syndrome and the making of scientific dictionaries! *Current Contents* (41):5-6, 10 October 1973.\*
4. ...., Everything you always wanted to know about ISI data bases but were afraid to ask. *Current Contents* (45):5-17, 7 November 1977.\*
5. ...., *The Permuterm Subject Index*: an autobiographical review. *Current Contents* (12):5-10, 21 March 1977.\*
6. **Greenberg D S.** Everything you always wanted to know about... *Washington Post* 31 January 1978, p. A19.
7. **Marx J L.** Calmodulin: a protein for all seasons. *Science* 208:274-6, 1980.
8. **Wade N.** Hybridomas: a potent new biotechnology. *Science* 208:692-3, 1980.
9. **Rembaum A & Dreyer W J.** Immunomicrospheres: reagents for cell labeling and separation. *Science* 208:364-8, 1980.
10. **Garfield E.** ISI adds "non-journal" material to the 1977 *Science Citation Index*. *Current Contents* (9):5-6, 28 February 1977.\*
11. ...., Five years of *Current Book Contents* and multi-authored book indexing. *Current Contents* (51):5-8, 18 December 1978.\*
12. ...., Announcing *Current Contents/Arts & Humanities*: in 1979 our *Current Contents* series will cover virtually every academic discipline. *Current Contents* (30):5-7, 24 July 1978.\*
13. ...., Introducing *Current Contents/Arts & Humanities*—a new service to keep you up-to-date. *Current Contents/Arts & Humanities* 1(1):5-7, 1 January 1979.
14. ...., ISI's new *Index to Scientific & Technical Proceedings* lets you know what went on at a conference even if you stayed at home. *Current Contents* (40):5-10, 3 October 1977.\*
15. ...., Introducing *Index to Social Sciences & Humanities Proceedings*—more help in locating and acquiring proceedings. *Current Contents* (33):5-10, 14 August 1978.\*
16. ...., Herpes simplex virus infections. Part 1. How widespread they are, and who is most threatened. *Current Contents* (25):5-11, 22 June 1981.
17. ...., Herpes simplex virus infections. Part 2. Sexually transmitted diseases without a cure. *Current Contents* (26):5-11, 29 June 1981.
18. ...., Leprosy: down but not out. *Current Contents* (37):5-12, 15 September 1980.
19. ...., All about ulcers, antacids, and how little we know. *Current Contents* (45):5-12, 10 November 1980.
20. ...., Introducing the *ISI Atlas of Science: biochemistry and molecular biology, 1978/80*. *Current Contents* (42):5-13, 19 October 1981.

\*Reprinted in: **Garfield E.** *Essays of an information scientist*. Philadelphia: ISI Press, 1980. 3 vols.