Reprinted from *Nature* 242(5396):307-309, 30 March 1973

The Synthetic Chemical Literature from 1960 to 1969

EUGENE GARFIELD, GABRIELLE S. REVESZ & JOANNE H. BATZIG

Institute for Scientific Information, 325 Chestnut Street, Philadelphia, Pennsylvania 19106

The literature of synthetic chemistry is growing at a rate of 8.7% a year—that is, doubling every 8.3 years. Data on 1.2 million compounds based on more than 128,000 abstracts published in *Index Chemicus* have been analysed. The articles abstracted contain 9.2 new compounds on average.

Current Abstracts of Chemistry and Index Chemicus is a weekly abstracting service covering published journal articles reporting primarily new chemical compounds, new reactions or new syntheses[1,2]. *Index Chemicus* (*IC*) was started in June 1960. Originally a monthly publication, *IC* became a biweekly in 1961 and a weekly in 1967. During the period of the study reported here, 1960 to 1969, all articles containing new inorganic as well as organic compounds were abstracted. The fundamental criteria for a "new" compound are, however, the author's claim of novelty, lack of a reference indicating previous publication of the compound, and therefore the presumption of novelty as for many intermediates.

This ten-year census updates an earlier five-year report[3]. It is not based on sampling. Rather, precise records have been kept for each journal. The number of articles abstracted and new compounds indexed for the calendar years 1960 to 1969 are shown in Table 1. The number of papers published each year almost tripled during this ten-year period. The first-year growth is atypical because the number of journals covered was then growing rapidly. Growth in subsequent years averaged 8.7%.

The number of compounds per abstract has stayed and remains remarkably steady at an average of 9.2 compounds an article.

By the end of 1969, records were abstracted of more than 1.2 million new compounds. By the end of the Second World War, there were approximately 1 million known chemical compounds in existence, so the new compounds reported between 1960 and 1969 are more than in the entire history of chemistry up to the end of the war. Between 1945 and 1959 another 1 million compounds were reported.

Table 2 shows the percentage of articles published in six of

the chief scientific languages—English, German, Russian, French, Japanese and Italian. Ukrainian is included under the heading Russian. "Others" include Spanish, Czech, Hungarian, Polish, Rumanian and so on.

English, as the leading scientific language, has steadily been gaining in strength from 50% of the total output in 1960 to more than 61% in 1969. German, which was formerly the leading scientific language of chemistry, diminished from 17% to 10% during the same period. This does not mean a decline in scientific output by German scientists. On the contrary, the growth of research elsewhere as well as the increased use of English has reduced the percentage, although there was a larger number of abstracts. The same is true of Japanese, which decreased from 7% in 1960 to 2.9% in 1969 during the most spectacular period of growth in the history of Japanese industry. There are many Japanese journals published today in English, and those journals published in Japanese have extensive English summaries.

Table 1 Compound/Abstract Ratios

Year	Abstracts	Compounds	Compounds/abstract	Increase in No. of abstracts % yr^{-1}
1960	7,581	70,408	9.3	
1961	9,167	87,496	9.5	
1962	9,899	94,172	9.5	7.4
1963	10,838	100,623	9.3	9.5
1964	12,084	110,230	9.1	11.5
1965	12,824	119,217	9.3	6.1
1966	13,943	128,830	9.3	8.7
1967	15,275	137,245	9.0	8.1
1968	17,830	167,340	9.3	10.1
1969	19,285	172,793	8.9	8.1
Totals	128,726	1,188,354	9.2 (average)	8.7 (average)

This same trend also prevails in the languages of other smaller countries—there are very few scientific articles published in Hebrew and Arabic, for example. Almost all papers from developing countries are published in English. A large percentage of Latin American papers appear in English-language journals.

This same trend is not true for articles published in the Soviet Union. Soviet scientists publish overwhelmingly in Russian with the exception of a relatively small number of articles which appear in multilingual international journals, such as *Tetrahedron*, *Tetrahedron Letters* or *The Collection of Czechoslovak Chemical Communications*. In these cases, Soviet scientists publish their work in English, German or French. Soviet chemical publishers are also still quite resistant to the idea of publishing English language summaries. In general, however, there has been practically no change in the percentage of Russian language publications during the past ten years. The average of 16%, although much lower than figures reported by *Chemical Abstracts* (18.4%)[4] for all fields of chemistry, is still high by comparison with Western countries. This is in part a consequence of the emphasis in the United States on physical organic rather than synthetic organic chemistry research.

Analysis by Journals

Table 3 shows the sixty top-ranking journals, as far as new chemical compounds are concerned, for 1969. It is interesting to note that eleven of the highest yielding twenty journals are published in English, and these accounted for 75% of the compounds. While many older well-established journals like the *Berichte*, *Annalen*, *Journal of the American Chemical*

Table 2 Language Distribution 1960–1969

Language	1960		1961		1962		1963		1964		1965		1966		1967		1968		1969	
	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C
English	49.9	49.4	53.0	53.3	52.1	50.9	53.8	52.5	55.5	51.5	55.2	54.6	56.0	54.0	56.9	55.5	61.3	59.7	61.4	60.2
German	16.9	18.7	16.4	18.8	16.2	18.3	16.5	19.1	15.9	18.7	16.4	17.4	15.3	17.4	13.8	15.4	11.8	13.5	10.0	12.6
Russian	14.7	13.1	16.0	12.6	16.4	13.3	16.2	13.4	16.8	14.7	16.4	14.4	16.4	13.9	16.4	14.0	15.4	12.5	17.4	13.8
French	7.5	8.6	5.2	6.1	5.8	7.3	6.0	7.2	4.8	6.9	5.9	7.1	5.5	7.6	6.2	8.3	5.7	7.4	5.9	7.2
Japanese	7.0	6.2	5.7	5.3	6.5	6.0	3.5	3.2	3.1	3.6	3.0	3.1	2.9	3.4	3.1	3.4	2.3	2.9	2.9	3.0
Italian	2.4	2.9	2.0	2.6	2.0	2.6	2.4	3.4	2.1	2.9	1.8	2.3	1.7	1.9	1.9	2.3	1.6	2.2	1.1	1.6
Others	1.6	1.1	1.7	1.3	1.9	1.6	1.6	1.2	1.8	1.7	1.3	1.2	1.9	1.9	1.6	1.3	1.9	1.8	1.5	1.2

A, abstracts; C, compounds.

Table 3 Number of New Compounds Reported in 1969

Rank	Journal title	No. of compounds	Rank	Journal title	No. of compounds	Rank	Journal title	No. of compounds
1	*J. Chem. Soc. (A, B, C)*	11,811	21	*Helv. Chim. Acta*	2,530	41	*Gazz. Chim. Ital.*	1,033
2	*J. Org. Chem.*	9,682	22	*Inorg. Chem.*	2,349	42	*Farmaco Ed. Sci.*	1,029
3	*J. Med. Chem.*	7,775	23	*J. Organometallic Chem.*	1,981	43	*Z. Anorg. Allgem. Chem.*	1,024
4	*J. Amer. Chem. Soc.*	7,638	24	*J. Heterocyclic Chem.*	1,954	44	*Roczniki Chem.*	1,001
5	*Bull. Soc. Chim. Fr.*	6,486	25	*J. Prakt. Chem.*	1,912	45	*Khim. Farm. Zh.*	997
6	*Tetrahedron Lett.*	6,272	26	*Austral. J. Chem.*	1,841	46	*Dokl. Akad. Nauk. SSSR*	961
7	*Chem. Ber.*	5,829	27	*J. Inorg. Nucl. Chem.*	1,723	47	*Arch. Pharmaz.*	945
8	*Tetrahedron*	5,525	28	*Arm. Khim. Zhur.*	1,575	48	*Arzneimittel-Forsch.*	900
9	*J. Chem. Soc. D Chem. Commun.*	5,415	29	*Zhur. Neorg. Khim.*	1,570	49	*Z. Naturforsch. B*	893
10	*Zh. Obshch. Khim.*	4,965	30	*J. Indian Chem. Soc.*	1,475	50	*Carbohyd. Res.*	769
11	*Zh. Org. Khim.*	4,398	31	*Acta Chem. Scand.*	1,407	51	*J. Chem. Soc. Jap. Pure*	751
12	*CR Acad. Sci. C.*	3,837	32	*Coll. Czech. Chem. Commun.*	1,407	52	*Ukr. Khim. Zhur.*	728
13	*Chem. Pharm. Bull. Jap.*	3,777	33	*Rec. Trav. Chim.*	1,259	53	*Biochemistry*	714
14	*Ann. Chem. Liebigs*	3,721	34	*Chim. Ther.*	1,151	54	*Rev. Roumaine Chim.*	684
15	*J. Pharm. Soc. Jap.*	3,195	35	*J. Pharm. Sci.*	1,123	55	*Khim. Prirodnykh. Soed.*	669
16	*Khim. Geterots. Soed.*	3,041	36	*Makromol. Chem.*	1,103	56	*Agr. Biol. Chem.*	662
17	*Canad. J. Chem.*	2,913	37	*J. Polymer Sci. A-1*	1,085	57	*Vysokomol. Soed. A*	660
18	*Indian J. Chem.*	2,806	38	*Angew. Chem.*	1,061	58	*J. Chromatog.*	646
19	*Bull. Chem. Soc. Jap.*	2,745	39	*J. Chem. Soc. Jap. Ind.*	1,060	59	*J. Agr. Food Chem.*	639
20	*Izvest Akad. Nauk. SSSR Khim.*	2,579	40	*Z. Chemie*	1,041	60	*Acta Chim. Acad. Sci. Hung.*	638
		104,410			30,607			16,343

Society, and *Helvetica* have remained fairly stable, others like *Bull. Soc. Chim. France* have doubled in content as have the Japanese journals published in English. But the newly-established journals like *Tetrahedron Letters* and *Journal of Medicinal Chemistry* have shown enormous growth—seven-fold and 4.5-fold, respectively.

Rapid communication journals have become very important. In 1960, the *Proceedings of the Chemical Society* contained 301 new compounds. By the time it changed its name to *Journal of the Chemical Society, Part D, Chemical Communications* in 1965, it contained 1,527 compounds a year, a five-fold increase, only to show a 3.6-fold increase by 1969 in which year it contained 5,415 compounds. In the same ten-year period, the yearly number of new compounds reported in *Tetrahedron Letters* increased from 876 to 6,272.

The output in the Soviet literature has grown considerably. Output for *Zhurnal Obshchei Khimii* grew steadily between 1960 and 1964, when it suddenly dropped off. This coincided with the publication of *Zhurnal Organicheskoi Khimii* and *Khimiya Geterotsiklicheskikh Soedinenii*. The combined output of these three journals increased by 100% during those years.

As shown in Table 3, analysis of journals by rank (in terms of yield of new compounds) shows that journals of rank 1 to 20 yield almost two-thirds of all compounds. Data are taken from the calendar year of 1969 but are also representative of other years.

Bradford's Law

According to Bradford's law[5] (illustrated in Fig. 1), one generally observes that a large part of the information in a specialized field can be obtained from a few top-ranking journals in that field. As the number of journals examined grows, the information gathered diminishes exponentially. In the case of synthetic chemistry, only seven journals were required to account for one-third of the compounds indexed in 1969. Only twenty-seven journals were required to cover two-thirds. More than 2,000 were needed to produce the last one-third. It is, however, more realistic to say that 158 journals account for the balance. In fact, 96% of the total output for 1969 was contained in only 100 journals. The dominance of a small key group of journals is much more dramatic in synthetic chemistry than that implied by a previous study[6], which reported that about 1,000 journals must be scanned to obtain 80% to 90% of the "chemical" literature reported in *Chemical Abstracts*. The *IC* census data correspond more closely to citation studies based on the *Science Citation Index* where it is found that the same small group of journals accounts for 90% or more of the citations appearing in chemical and other journals[7-9].

To support this finding to our own satisfaction, a separate study was conducted from 1967 to 1969 to make sure that no important chemical journals was overlooked. Article titles from 2,000 additional journals were searched by computer using ISI's Automatic Subject Citation Alert (ASCA). The project was discontinued when it became clear that these 2,000 journals only contributed 1% of the output obtained from less than 200 journals. As these journals were all covered in the *Science Citation Index*, the probability that pertinent information would be lost in a comprehensive literature search is very low indeed. Any cost-sensitive approach to information management must set some upper limit on the scope of any service[10].

Duplication for Searchers

The chief conclusions of our study are as follows. First, the synthetic chemical literature has been growing steadily at a rate of approximately 8.7% a year; second, English is in-

creasingly the leading language of chemical communication and about two-thirds of the newly prepared compounds are reported in English; third, one-third of all new chemical compounds are published in seven primary journals, half in fifteen, two-thirds in twenty-five; only 100 are needed to cover 96%.

As far as new synthetic chemistry is concerned, a cost effective service is provided by examining a carefully chosen

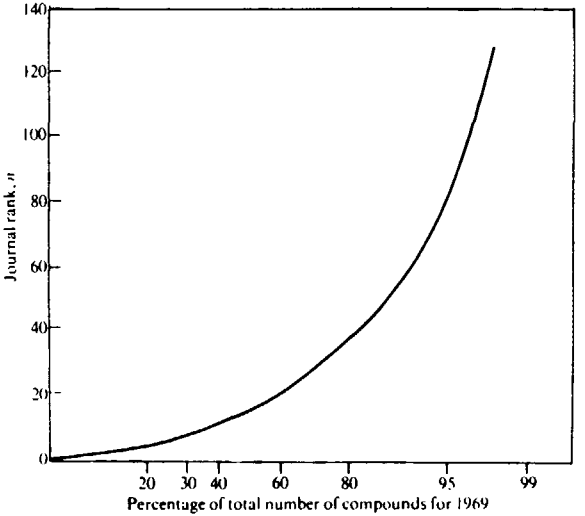

Fig. 1 Percentage of total number of compounds reported in 1969 for journals of rank $< n$. For $n = 1$–20, the number of compounds is 104,361; for 21–40, 30,656; for 41–60, 16,343; for 61–80, 7,154; for 81–100, 5,238; for 101–183, 7,041.

and monitored list of 100 to 200 key journals. Although there may be other justifications for monitoring 12,000 journals at *Chemical Abstracts*[4], it certainly cannot be justified for the needs of synthetic chemistry.

Limitations

It may be argued that additional new compounds are found as a consequence of abstracting for other branches of chemistry, but there are economic limitations involved when one considers the exponential increase in computer and other costs involved if one must check all compounds, new and old, through a large-scale registry. Furthermore, the more intensive line by line examination and indexing of the more significant articles in synthetic chemistry, by contrast with indexing from abstracts, produces a number of not otherwise reported compounds, including thousands of intermediates.

[1] Elias, A. W., Revesz, G. S., and Foeman, G. H., *J. Chem. Doc.*, **8**, 74 (1968).
[2] Revesz, G. S., and Warner, A., *J. Chem. Doc.*, **9**, 106 (1969).
[3] Garfield, E., and Foeman, G., *American Chemical Society 148th Meeting, Abstract*, No. G31 (September 1964).
[4] Baker, D. B., *Chem. Eng. News*, **49** (28), 36 (1971).
[5] Bradford, S. C., *Documentation* (Crosby Lockwood, London, 1948).
[6] Wood, J. L., *The Parameters of Document Acquisition at Chemical Abstracts Service, American University Eighth Annual Institute of Information Storage and Retrieval* (Washington, DC, February 14–17, 1966).
[7] Garfield, E., *Science* 178:471 (1972)
[8] Garfield, E., *Current Contents*, No. 5, 6 (February 2, 1972).
[9] Garfield, E., *Current Contents*, No. 9, 5 (March 1, 1972).
[10] Weinstock, M., *Nature*, **233**, 434 (1971).