E. de Grolier

Problems in Scientific Communication

Introduction

The intent of this paper is to relate *scientific communication* to its historical and cultural background. Both *communication* and *science* are taken here as somewhat comprehensive terms, including personal, direct contacts as well as documentary communication.

By way of introduction, I should say that, gauging communication efficiency by the speed of diffusion of scientific and technical discoveries, we have indeed made very good progress during that part of scientific communication history which we have witnessed. Contrasted with the very slow diffusion of information in prehistory, for instance in ancient agriculture, cattle-breeding, iron technology or in the development of the alphabet, and even in modern paper manufacture and printing, the speed of diffusion of scientific culture has increased tremendously, and that diffusion has also become much more universal. This is naturally only a part of the general process that the French historian Daniel Halévy called the *acceleration of history*.

But this is not to say that our systems of scientific communication are already perfect. On the contrary, there is room for considerable improvement. For example, even today we may cite important discoveries which take a very long time to be properly applied and to penetrate into the common practice of science or technology all over the world. One need only recall certain discoveries of science or technology made by Leonardo da Vinci, which were until now embedded in the mass of his manuscripts. From time to time one of these is rediscovered, such as da Vinci's fine machine for making pins. In our own field, there is the famous peek-a-boo, discovered by an American in 1915 and rediscovered two or three times in England, in France, and in the United States.

Then there is the inefficiency in terms of the disproportionately large amount of time consumed by scientists and technicians in searching for information, as against the relatively short time left free for creative research (one of the results of the study currently being made by Ackoff at Case Institute). One unfortunate solution consists in neglecting useful information which might save many fruitless efforts. Witness, for example, the percentage of so-called inventions or discoveries which are nothing more than *vieux-neuf* — "old-new." There are numerous world-wide examples in patent offices of "inventions" that are not inventions at all.

Finally, there is the relative inefficiency of our communication mechanism for general diffusion of scientific and especially technical knowledge among laymen—the relative failure of the so-called popularization of science, on the one hand, and of technical information for the majority of enterprises on the other, the latter being revealed by the recent European Productivity Agency inquiry in Europe.

In what follows, we shall examine, first, barriers to scientific communication, and, next, means for removing these barriers — aids that could be found for improving scientific communication.

Barriers to scientific communication

We will discuss eight barriers, though perhaps there are many more. Barriers (1) to (6) are barriers outside of science itself, while barriers (7) and (8) stem from the development of science.

• 1. *Distance*

By the barrier of distance, I mean not only the physical distance between scientists, but also that distance imposed by the mores of different lands—a sort of cultural isolationism. As a small example, I could cite a certain method, used in Central Italy more than a century or so ago for fighting soil erosion, which has not yet penetrated into France, although specialists believe it could be adapted to many areas there.

• 2. *Nationalism*

Physical or psychological barriers to scientific communication are frequently erected by nationalistic behavior. An overly-zealous customs administration, for example, may hamper or even bar the entry of "foreign" films or punched cards or whatever on trivial grounds. I would like to cite one tragi-comical case: a technical film on steel manufacture which was rejected by the customs authorities of one European country who feared that it would be of the Folies Bergères type. Or one nation may assume a mental attitude that scoffs at what is being done in various areas of research by "foreigners"—in German physics, for instance, or in Soviet biology, and so on; and this attitude is very commonly adopted more or less unconsciously, in all countries, including mine and yours.

For example, I recently saw a certain article on classification (which is my specialty) in an American periodical with not a single reference to work done on the continent of Europe.

• 3. *Secrecy and censorship*

These are two very old evils, dating at least from the Egyptian priests and seemingly everlasting, recurrent now after a temporary eclipse during that sort of Golden Age of liberalism that was the Nineteenth Century.

Some periodicals also practice a "concealed" censorship by declining to publish such information from foreign contributors as might conceivably compromise a national advertiser.

• 4. *Prejudice against science*

This evil is probably less acute now than, say, two centuries ago, but it has certainly not yet disappeared, even in our countries, not to mention others less developed than ours. Witness the disproportion of graduates in scientific fields to those in the so-called humanistic fields. And traditionalism is still very much a fact.

• 5. *Ignorance*

The inability of a majority of people to understand scientific method and scientific language is of course disappearing, too, but rather slowly. It would be difficult indeed to speak of scientific communication to a population of illiterates; but although illiteracy is no longer general, what about the semiliteracy of a population which has an average mental age of ten or twelve, which is, I think, the average mental age in France or the United States?

• 6. *The Tower of Babel of language*

This barrier is even more serious. Because there is no longer a dominant language for scientific communication, the situation is deteriorating rather than improving. It is a big mistake, I think, to believe that any one natural language can be the world language of science. Would it be English? Let us remember the instance of Rivarol, who wrote in France about the universality of the French language at precisely the time when this universality came nearest to being a fact, but thereafter became less and less a fact. Now there is a growth of scientific literature in Russian, in Japanese, and soon in Chinese, not to mention the Bantu languages which are coming along. Urquhart made the rather disagreeable prophecy that at the end of this century one-half of world scientific literature would be in Chinese. If I live that long, I shall not be able to understand half of the scientific literature—unless, of course, I decide to learn Chinese. Might the world language of science be Greek? This was the thought of Hogben, another British scientist and specialist in general linguistics, and perhaps it is more in keeping with the facts (though not for the Chinese, anyway).

These, then, are the barriers outside of science. Now let us look at two barriers coming from the development of science itself.

• 7. *Specialization*

Specialization is a very good thing in itself, and cannot be avoided, but it is nevertheless a very effective barrier to scientific communication. It impedes cross-fertilization of one scientific domain by another, and it hinders the application by one category of specialists of useful methods, instruments, et cetera, devised by another. It is a second Tower of Babel, a confusion of tongues within one and the same tongue. And it is a very serious fact, considering that the most important advances in science have always come from the fringes of, or frontiers between, different disciplines. Or consider the success of men trained in one discipline and working in another, like the great Pasteur, who was not a medical man but who, however, probably did more for medicine than any other in his era.

• 8. *Scientific literature*

This formidable barrier is created by the sheer bulk of existing scientific document production. For example, of 20 million documents analyzed by *Chemical Abstracts*, 40% were produced during the last 10 years. Then there is the rapid growth of the scientific community itself. Urquhart speculated that at the end of this century there would be twenty or thirty million instead of one million scientists in the world. Obviously, the means of communication that are efficient and practical for a determined amount of documents and a limited number of scientists or technicians are no longer efficient and practical when these are multiplied, say, by ten or twenty or a hundred or a thousand.

At one time private letters were quite sufficient, as, for example, those written by Descartes, Huygens or Newton. After that there were the periodicals, and then technical reports. At one time a personal index or a few notes on a piece of paper sufficed. Following that came bibliographies and abstracts.

This entire system is now collapsing under its own weight. Whether we wish it or not, it is absolutely necessary to find new ways of publication, retrieval, selection, and utilization of scientific information, in a word, of scientific communication. I use the term *entire system* deliberately, because all of these form a system, and when you change a part of the system, you must change all of it.

There is certainly a very real urgency about the need for a change. We speak of the need for a "brain-pool," a somewhat vague expression the best meaning of which, in my opinion, is a drastic improvement in scientific communication. Or, to take up the question put by the late German librarian and documentalist, Hugo Krüss, the problem is how to master, to dominate, knowledge.

Aids for improving scientific communication

What means are available for removing the barriers discussed thus far? By *means*, I intend not only mechanical aids but also new types of organization. Consider barriers (1) through (5). Concerning distance — anything that

increases the speed of transportation is very important for scientific communication. A few decades ago air travel had not advanced to its present stage, which permits my being here today. As for radio, television, teletypes, and the other means of long-distance communication, new improvements are constantly in the making.

Next, we have many methods that facilitate communications between different nations or cultural areas: the sending abroad of trainees, missions, the exchange of professors or researchers, exchange of publications, technical *assistance,* or, better, technical *cooperation*. There are questions here of politics and economics, and therefore partly outside the scope of this conference and beyond the influence of scientists themselves. But scientists do have their say concerning such important and vital matters as secrecy. I commented in France recently that your great scientist, Oppenheimer, had made some very keen remarks on this subject.

Some progress has been made by UNESCO and by progressive nations like the United States. I would cite an experiment sponsored by the French Government which I am now conducting. Our aim is to gain experience in establishing centers to facilitate the exchange of scientific and technical information and personnel between France and other nations. The first of these was created a few months ago for relations with Italy. But there are difficulties not easily overcome. The system of national patents, for example, has served the development of industry very well. But I think now this same system is so overly protective that in many instances it becomes a real hindrance to effective scientific communication.

The barriers of prejudice and ignorance are problems of education, broadly conceived; for the most part they fall outside the sphere of action of scientists themselves. But not completely. I think the scientists have a certain responsibility for popularizing the results of their research to the layman. And it is also the responsibility of librarians and of technical writers (or rewriters). The United States has many institutions which are real models in that field, from the Smithsonian Institution to the Science Service in Washington. But there are serious problems to investigate—possibly with scientific methods —such as the problems of rewriting, of increasing so-called "readability," and so on.

The problem of language would require an entire conference like this for adequate treatment. In fact we hope UNESCO can organize such a conference in the not-too-distant future. The gravity of the language barrier has been commented upon by men such as Dr. Oppenheimer, and a very good book on the subject was recently written by Dr. J. E. Holstrom. Further recognition of the problem is evidenced by the existence of "IACDocTerPAS": the International Advisory Council for Documentation and Terminology in Pure and Applied Science (whose abbreviated name, by the way, reminds me more of some prehistoric animal than of a scientific body).

Language is also partly a problem of education—education of the general public and of the scientists themselves. For instance, in the Netherlands, Switzerland, or Denmark, there are scientists who have fluency in three or four languages rather than just one. Further, the problem is partly one of techniques and machinery, and here we have the promising beginning of mechanization. But it is also partly a problem of organization, of translation pools, possibly international, possibly translation *en masse* —mass translation from cover to cover. There is in this field a real need for a well-planned program on a large scale, such as was suggested in England by Vickery. All this is based on the hypothesis that the present situation of multiplicity of languages for transmitting scientific information is taken for granted, and left unchanged. We could possibly think of more radical solutions: first of all, the possibility of an international, artificial, or, more exactly, rationalized language for science. The fact is that all attempts made until now in that direction have proved inadequate. But there have been some promising attempts, such as those by the Australian, C. K. Bliss, and by the American, S. C. Dodd.

In this respect, we might compare linguistics with the history of aeronautics. In the first stage, there were *a priori* plans insufficiently supported by the actual progress of knowledge and techniques—in aeronautics, the design of Leonardo da Vinci for aircraft; in linguistics, the beautiful though unrealizable plans of Leibniz in Germany or Wilkins in England. The second stage was the imitation of nature—flying men died because their wings were artificial and inferior to those of birds; and Esperanto, Interlingua, Interglossa, and the like, also died, more or less, because they were inadequate. The third stage in both cases was the research for new solutions, not *imitating* nature, but *inspired by* nature, and original. Finally, just as there were two directions in aeronautics, craft lighter and heavier than air, there are two directions in international language, pasigraphy and pasilaly.*

The powerful support for progress coming from the field of coding is an example of mutual fertilization of one field by another. There are aids, too, in the form of linguistic research, especially "structural" linguistics, from semantics, from the study of symbolism in general, from logic and mathematical logic. There is a real and urgent need combining all this into a unified research endeavor.

Next comes the problem of specialization. The means for removing this barrier to scientific communication involve the organization of research itself; for instance, better organization of teams for research. We have very few efficient groups consisting of specialists representing fields different from each other—in operations research, for instance. In France there is a Groupe Français des Argiles (a French group for research on clay) which is composed of fifteen specialists from about as many different disciplines, which has done very good work just

**Pasigraphy:* Any artificial written language designed for universal use. *Pasilaly:* A universal (spoken) language. (Webster's New International Dictionary, Second Edition, Volume III, G. & C. Merriam Company, Springfield, Mass. 1954.)

because of the diversity of backgrounds. There should certainly be created an organization for counterbalancing the effects of overorganization and overspecialization. We should also reinforce our whole system for compiling bibliographies, digests, annual reviews of progress, as well as special kinds of abstracts made for other specialists. There is some promise in that direction in the researches into auto-abstracting techniques conducted by H. P. Luhn of IBM.

The problem of establishing connecting links between different documentation centers and/or general documentation centers was attacked first, I believe, by France in 1939 with the creation of the Documentation Center of the National Center for Scientific Research. The VINTI in Russia adopted the idea of the CNRS, but with means a hundred times superior. A similar proposal for America has been made quite recently by Stanford University. But there are many lines of attack. We could think of taking the results of research obtained from all disciplines, on the methods or instruments employed—here we may refer, for example, to the work of Wildhack at the National Bureau of Standards.

Concerning the linguistic problems arising from specialization, there is the suggestion of the Frenchman, Robert L'Hermite, that scientific literature be systematically rewritten for *other* scientists not specializing in the same field. But the ultimate problem concerns the overall reform of scientific language itself.

The problems arising from the sheer bulk of documents comprise the last, and probably the most important, barrier. The first line of attack is on the front of publication; more precisely, on two fronts: repetition, or the inflation of documents on the one hand, and the scarcity or nonexistence of certain documents about a given subject on the other. The Spanish philosopher Ortega y Gasset commented years ago that there are at the same time too many and too few books. That is very true, and it implies problems of planning, of choice, and of selection. Certain tendencies perhaps could or must be counteracted: for instance the problem of university inflation of papers, based on the desire for academic promotion. It is a very real problem in some places, in that it means more money for a professor if he writes, say, ten reports a year instead of one, even if there is material for only one. This proliferation is also due to some policies of scientific and technical periodicals, as well as of certain scientific and technical conferences. Rather than this policy, perhaps there should be a policing of documentation, as for traffic. And I think this policing could be done by scientists and documentalists themselves, much better than if the powers were delegated to an outside police.

Next, there is the problem of so-called *bibliographical control,* which arises from the failure of the traditional bibliography. There are omissions pointed out, for instance, by the Coblans and Kyle's reports to UNESCO, on the lack of control for new forms, such as technical reports or films. There is a need for balance between bibliographical control and supply of actual documents which brought about such plans as the British Central Lending Library for Science and Technology; microfilm service of the French Center for Scientific Research; or the Farmington Plan here. All very well, but I think more, much more, is needed now.

We hope there will be adequate discussion of these problems at the November International Conference on Scientific Information in Washington. For example, should there be a centralized organization for documentation, or not? What of abstracting systems? The promise of so-called auto-abstracting as set forth by Mr. Luhn's research into this possibility could replace abstracters by machines. What of surveys? There is much to be said for the practice of what is called in France *notes de mise au point,* for instance, at the Center for Research on Coal in France (which analyzes and subsequently synthesizes all documents pertinent to a given question). This is done on special request of the research team, at the very beginning of the research by a special information officer, but in cooperation with the research men themselves. There are also the *traités de mise à jour,* that is, treatises digesting a whole field, such as that made by the same L'Hermite at the RILEM, Réunion Internationale des Laboratoires d'Essais des Matériaux (International Group of Laboratories for Testing Materials).

We must distinguish different "levels of analysis" as Vernon Clapp once said. All levels are useful. It is very useful to have good directories of documentation centers, good and currently revised union catalogs, but we need above all "depth documentation," so called by Ranganathan. The goal is, after all, the information contained in the documents, and these are only means to obtain that information. And this introduces us into the whole, almost unlimited, field of *information retrieval,* and *mechanical selection of documents.*

We are concerned here with two different things: first, machinery; and second, language for that machinery—or, to put it another way, memories and coding.

Memories

Concerning memories, I think that perhaps too much propaganda has been made for individual efforts and particular gadgets. Each inventor is naturally convinced of the indubitable superiority of his own system. But I should say that certain of these systems are overpraised and overadvertised — for instance, marginal punched cards. I had a somewhat disagreeable experience with the use of marginal punched cards in a project for the French Operations Research Society — the system collapsed at just five thousand documents. Four thousand nine hundred and ninety-nine were good; but when we had five thousand it was really bad. And I think similar systems are also heading up a dead-end street. I personally think that this is the case with normal punched cards in this particular field of information retrieval, even if IBM has made its fortune with them in other applications (and I think that IBM people think so too, after all).

The same seems to hold true for solutions like the Samain's Filmorex or even the Kodak Minicard, which

involve the processing of both the documents themselves and their indexes as a unit, on the same support, and with the same machinery. This is not, in my opinion, an economic solution. Then where do we look for the most promising system?

If we exclude those systems which are good for some organizations or some processes but in very limited fields like "Synoptic" or Kardex-type cards, there probably remain but two main categories of machine systems: (1) the so-called "peek-a-boo" systems, and (2) systems requiring the adaptation of computers.

Someone has explained to me that the "peek-a-boo" system can best be described by the term "in one ear and out the other," that is, a simple input-output system. In their present form, these systems are too manual, too slow in operation, too restricted as to the number of documents they can practically control, and too susceptible to errors.

French work along these lines involved a series of machines studied by Gerard Cordonnier, and consisting of machines for reproduction of his "Selecto" cards, their recapitulation or summing-up, the automatic recording of characteristics, a selector giving mechanically prepared lists of call-numbers of the documents, and a print processor to automatically deliver microfilms of those documents. Unfortunately, the man who was in charge of designing the machines destroyed them in a fit of insanity, but I sincerely hope that another French engineer will yet complete the work.

Other possibilities include the patents submitted four or five years ago to the U. S. Patent Office by the Régie Nationale des Usines Renault and which I think have not yet been examined. Or, there is the idea of Batten in England for the use of magnetic tapes, each one recording the call numbers of documents corresponding to a given characteristic to be searched simultaneously. This or another method could be coupled with an organization dividing the whole field of knowledge among a certain number (not too many) of specialized documentation centers, adopting the same method and the same codification, under control of a universal coordination center. This was Cordonnier's proposal made at Dorking, England, last year and now seriously considered by our Centre National de la Recherche Scientifique. The plan at first seems utopian in concept, but similar reorganizations are going on in certain other fields of human activity, and I believe an objective study in this area will prove profitable.

The second category of machine systems is that requiring the adaptation of computers. I use the word *adaptation* advisedly, for in this respect computers badly need to be adapted. I think that a lot of criticism by Ralph Shaw and others directed against the use of computers for information retrieval is fallacious in the sense that these criticisms are appropriate to existing computers but not to those designed especially for use as information-retrieval machines — and not for quite another use, calculation.

We must have very large memories with very rapid access. These are two conditions which seem at first to be irreconcilable, but in fact there are probably means to reconcile them. I allude to the work done by IBM's Gilbert King or M. M. Astrahan; or to the Russian claim (by Gutenmacher in a recent issue of *Akad. Nauk. Izvestia*) of a machine with a large memory and very rapid access; or to work in progress in France on Bull machines by Dr. Brygoo, and in the French Atomic Energy Commission by Braffort, and the work of F. Raymond for the SEA, the Société d'Electronique et d'Automatisme.

I also want to call attention to the progress made in the practical and easy conversion of symbolism — one sort of codification — to another. I am thinking of apparatus which converts printed or written characters into machine language, as produced for example, by the Intelligent Machines Company in Virginia and the Solartron Electronics Group in Surrey, England; or of those converting machine language into ordinary alphabetic writing, like the Compositron of RCA or the Numérographe of the Société d'Electronique et d'Automatisme; or of the other machines doing the same kind of work with the spoken language like Sonographe or Vocoder, which produce sounds approximating language, using magnetic tape—although I must say they produce more noise than language, at least for the present.

Mr. Luhn pointed out to me the very important fact that there already exists a great mass of documents, written in a common, ordinary language, which are automatically converted into machine language: for example, all the production of teletypewriters and teletypesetters, or of monotype machines. Here is a great potential source of material for feeding the machines working at information retrieval, and measures are now being taken to save this material directly appropriate to machine processing of information.

I might add that the technicians can probably gear the electronic machines to do all we would like them to do, but first they must know what we want—for what purpose, for what volume of information, for what kind of work and problems.

Coding

The problem of coding must be attacked with much greater intellectual effort, and I believe industry and government must be prepared to compensate such greater effort with comparably greater financial reward.

Creating a general machine language is, in my opinion, a necessity which simply cannot be eluded. This could also be the link to join the different translating machines working with natural languages. There are many points to be very seriously considered in designing this machine language, and many scientific disciplines must be brought together for this common work. Classification is one of these, linguistics is another, logic still another, and I can allude to some names of persons whom I know and who work along these lines: Ranganathan in India, Vickery or Foskett in England, Belevitch in Belgium, Scheele in Germany, Rennes, Pagès, Durocq in France and, if this

kind of auto-citation is permitted, myself, with the study of general categories to be applied to many different fields of knowledge, such as logic, morphology, space, time, and so on, under contract with UNESCO.

In America you have a good team at the Patent Office with Andrews, Newman and others; I have already cited the names of Bliss and Dodd, and there are probably many others unknown to me. Incidentally, one of the first to evoke the possibility of such a language was the late George Orwell in his novel *1984*—a kind of artificial language which reminds me curiously of the Andrews language.

One can envision going a step further, and in fact we may prophesy that we are to go further sooner, perhaps, than we can now imagine. One may conceive of machines capable not merely of retrieving information but of elaborating and transforming information into new knowledge. I am not at all in agreement with those who thrust aside such ideas as utopian or even simply as madness. In fact, there are men like John M. Maguire or Vincent P. Biunno, Lucien René Mehl and his fellow workers at the Conseil d'Etat in France, who are thinking very seriously of the possibility of mechanizing the work of judges (and perhaps to do their work better than they do it now). Dr. François Paycha in Montpellier has already mechanized diagnosis in ophthalmology and is preparing the way for mechanization of prescriptions—of therapeutics —in the same medical specialty. He thinks that his machine "doctor" would be much more reliable than the average practitioner.

It was in *The Revolt of the Masses* that the Spanish philosopher, Ortega y Gasset, wrote that it was really rather easy to foretell future events, and that there were numerous examples of men capable of describing in advance the course of events which would later profoundly modify the face of human societies. I think this is perfectly true of scientific communication, and I think that soon the dreams of Leibniz, Descartes, or Wilkins will be finally realized. The engineers and scholars who are working in this most important field are contributing to what shall be a great and profound revolution which, transforming the aspect of scientific communication, shall also change—for the better—the destiny of the human race.

Received May 27, 1958