Contents

Chapter 4

Chapter 3

The Language of Science

Terms in Science

In this section I make an attempt to consider some metalinguistic problems of the language of science. First of all, I should like to answer the question: How do scientific terms appear and evolve; why and in what way do we comprehend them? Strange as it may seem, it is not very easy to speak about this. Since our school days we are taught the concepts of some indubitable rigor and precision of judgments in science. Hence, it seems that in science we should have realized that the terms (at least, partly) do not come into being in the same way as new words do in our everyday language. So far, widespread opinion has been that the terms should be defined in the rigorous logico-linguistic sense of the word. That would mean that, when a new theoretical term is introduced, it is once and for all time ascribed a strict sense content expressed in a defining phrase; but this is hardly the case. Carnap's elegant concept of the semantics of scientific language was built not inductively, as a theory trying to comprehend and systematize actually observed facts, but deductively, as some idealized, logically perfect system—a program for the future. If somebody still feels like tracing in what way it is possible to approach the analysis of scientific terminology from the formally logical standpoint, I shall send him to the book by Hutten (1956), *The Language of Modern Physics*, which is highly readable and has already been mentioned above.

Another purely empirical approach seems to be much more interesting. This is a many-sided, theoretically unbiased analysis of the whole semantic diversity of scientific terminology. The recent book by Achinstein (1968) is written from this standpoint. If one takes this road, many interesting observations may be made, but even here it can hardly be hoped that some sufficiently general theory can be built which would explain the whole diversity of the phenomena observed.

I shall illustrate the complexity of the mechanism of forming scientific concepts by several examples. First of all, Achinstein draws our attention to the fact that some terms such as "copper," "metal," "metalloid," "brass," "semiconductor," "bronze," "fusion," etc., may be regarded as certain taxons—the elements obtained as a result of classification performed upon some multidimensional space of attributes. These attributes may not be necessary: the metal mercury is a fluid under normal temperature, though solidity is one of the indications of a metal. They may also be insufficient: one of the properties of copper, its melting temperature of 1,082°C, is not sufficient, because some other alloys have the same melting point. With scientific development, new attributes appear and the old classification schemes become insufficient. Thus, the traditional division of chemical elements into metals and metalloids also proved insufficient after the emergence of a clear-cut concept of semiconductors. True, the latter is already connected with the physical properties of the substance but not with its chemical ones. But in fact we do not know for sure which properties are chemical and which are physical. The demarcation between physics and chemistry has become blurred, though clear-cut distinction of these terms is of great pragmatic significance. In the editorial boards of the "abstracts" journals there are continuous and unsolvable arguments as to the headline under which this or that article should be published, either under "chemistry" or "physics," and when a new leading scientist, a chemist or a physicist, enters the editorial board, the headings of the respective issues are inevitably changed.

In the process of constructing the system of scientific terminology, we face the same difficulties as in ordinary statistical problems of multidimensional classification since scientific terms can be regarded as taxons, i.e., units of classification given over fuzzy fields of meanings. In statistical classification methods, there are a variety of techniques leading to essentially different results. For each technique not only is a classifying procedure given, but there is no stopping rule; in any case, there are no criteria which would permit one to insist that further division does not lead to obtaining actually new taxons. Every result of classification may be essentially changed, if the metrics of the space of variables is changed. In our case, it will mean the equivalent of ascribing various weights to certain properties. We should also keep in mind that taxonomy is a purely semantic problem and not in the least ontological. There is no use raising the question about whether there are some realities in the world of things corresponding to taxons. In any case, the criteria for this cannot be established.

No doubt, in scientific language there are some terms which may be clearly defined, say, "Bohr's atom," "the black body," "double-atomic molecule"; these are the examples taken from Achinstein's book. But it often happens that the definitions which look quite respectable from a logical standpoint prove insufficient; they cover too wide a range of things. This happened to the term "document." It could be defined as follows: "*a document is any material carrier on which certain information expressed in a language is fixed.*" If such a definition is accepted, then the fence with the joke written on it by children immediately turns into a document, though it does not correspond to our intuitive notion of a document. We may attempt to save the situation by introducing the supplementary series of *operational* characteristics, and declaring that not any information carrier is to be called a document, but only the one which bears certain numerable functions. In many cases the introduction of operational characteristics helps considerably. But in our example with the word "document," it makes everything look more anecdotal: after the well-known apt expression, it turns out that the elephant in the zoo is a document, and all the other elephants are not. But at the same time, we know for sure what a document is. Moreover, there exists a special scientific branch called "documentation science"; certain journals are published whose title contains the word "document"; and in modern soceties dozens of thousands of people are engaged in "documentalistics" as a scientific discipline.

Evidently there is no sense in a further consideration of all possible ways in which terms emerge in science. The above examples seem sufficient to illustrate the difficulties encountered by scientists when they try to outline the boundaries of a term.

Now we shall go into more detail with another question, that of the analysis of certain phenomena which are specific for the language of science as a whole. Here we shall speak of the profound connection of terms with theoretical notions in science, of their code-wise character, of the role metaphors play in the generation of scientific words and of the possibility of using rather abstract notions with their ambiguous interpretation in the language of concrete representations.

The connection of terms with theory. Terms in science are closely connected with its theoretical concepts (Feyerabend, 1962). On the surface, many terms seem to be no more than names of some objects or phenomena. For example, the Raman effect would seem to be a name of some physically observed phenomenon. In fact, it is neither through the indication of what it denotes nor through some semantic definition that the meaning of this term becomes clear, but through the understanding of the theory of this phenomenon. The same is true of such terms as "atom," "electron," etc. A pupil in a secondary school attaches a different meaning to these words than a physicist does. The meaning of the words changes with time together with the development of our scientific concepts. In any case, the meaning we ascribe to the word "atom" differs considerably both from that ascribed to it by the ancient Greeks and from that used at the beginning of our century. But there is one especially interesting thing: it is possible that in science several concurrent hypotheses exist simultaneously which use the same terms but in a different sense. There can also exist some theories, one being above the other or one including the other, e.g., the relativity theory and classical mechanics, and they may both use the same terms in essentially different senses. Both classical mechanics and relativity theory make use of such terms as "mass" and "length," but they are interpreted differently. When we speak about space in physics, we may mean the space of both Euclidian and non-Euclidian geometry. And what is exceedingly surprising is that, as a rule, in science all this causes no trouble of the kind it does in philosophy. Being the adherents of different theories, scientists may use in an argument the same words in different meanings. From the standpoint of logical semantics, this does not seem possible (Achinstein, 1969). Actually, it becomes feasible when the word is associated with a distribution function of the meaningful content, given in scientific terminology by a scientific concept. Different scientific concepts will lead to different, though correlated, distribution functions of the meaningful content. From the semantic standpoint, a scientific discussion may often be regarded as a procedure aimed at the improvement of the correlation of the prior distribution functions of the meaningful content of the term.

The metaphorical structure of the language of science. If, reading a scientific text, we stop for a moment and ponder the character of terms in our field of vision, we shall find that they are metaphorical. We have become so used to metaphors in our scientific language that we do not even notice it. We keep coming across such word combinations as "course of time," "the field of force," "temperature field," "the logic of experiment," "the memory of a computer," which allow us to express new notions with the help of rather unusual combinations of old, well-known, and familiar expressions. Recognizing the right of metaphors to existence in scientific language, scientists have permitted rather different senses for old terms with the emergence of these new theoretical conceptions. In science, theories are continuously changing, but the change does not cause a waterfall of new words. The new phenomena are interpreted through the old, familiar ones, through the old words for which the prior distribution function of meaning is slightly, but continuously, changed. Something remains unchanged but becomes of less importance; something new appears, entirely different from, and to a certain extent contradictory to, the former meaning of the word. Now the role of metaphors in the language of science is evident; it is alluded to in a most elegant way in MacCormac's work (1971). It is his example with reference to Feyerabend (1965) that is especially interesting. There he speaks of the term "force," one of the fundamental terms in physics. Emerging on the basis of the notion of human force, it has undergone a long history starting from the neo-Platonic philosophy via Kepler and Newton up to modern physics without being strictly defined; always it has remained at the metaphorical level. And the especially intriguing fact is that, obeying some unconscious inner pressure, entirely new terms are introduced into science as metaphors. In mathematics there recently appeared such metaphorical terms as "group," "bodies," "rings," "regression," and "regression analysis" (literally, regression means "backward movement," "reversion"), and "mathematical expectation"; eventually, all the terms received strict definitions. I shall dwell at more length on the meaning of the last term.

In ordinary speech the word "expectation" is modal and is used when something is expected; i.e., it may happen, but it may not happen as well. We can expect the weather to be nice tomorrow or our friend to arrive, but if today is Friday we shall never say that tomorrow is expected to be Saturday. And when we speak of the mathematical expectation of a random value, we mean its average which will necessarily take place, if we carry out averaging on the indefinitely large number of observations forming the so-called general population. The modality of the word "expectation" has undergone changes here, and this change is not logically conditioned by adding one more word. It is just to a certain, fixed combination of two logically incompatible words that we assign a particular sense.

"A new confrontation of words must create either strain or absurdity. If a metaphor does not provoke thought, then it appears as a symbol rather than a metaphor" (MacCormac, 1971). The metaphorical structure of the language makes it not only polymorphic, but also strained. Above, I have already spoken of the observation by Podgoretskii and Smorodinskii (1969) to the effect that a new axiomatic basis in physics emerged after the revelation of hidden contradictions in previously published papers. For a while these contradictions remained unnoticed, evidently, precisely by force of the metaphorical structure of the language, though, of course, this question needs further investigation. But it is noteworthy that, when introducing new words, scientists often yield rather to psychological influences than to logical ones. This can be easily explained. The transmission of thought is carried out on a logical level, but its perception is greatly influenced by some psychological factors which are not entirely understood. An idea is perceived more readily if it is shocking and requires an intellectual effort. A good scientific paper ought to be a bit incomprehensible; there is nothing like some reticence to express peculiarities of ideas. The papers which are too comprehensible seem childish. Incomprehensibility is most often created by some deliberate linguistic structure, which often becomes lost when the work is translated into another language: we sometimes cannot recognize our papers after translating—so dull do they become. Use of the metaphorical structure of language is only one of the techniques used to create intellectual strain. The creators of Zen culture apparently understood this psychological peculiarity of the mechanism of perception of complicated concepts especially well, and they widened its use to the extremes, introducing koans as special illogical forms of thought transmission.

Polymorphism of scientific terms. Scientific terms have a more polymorphic character than the words of ordinary language. This is only natural: they contain more meaningful content than the words of our ordinary language. We may give innumerable examples to illustrate polysemy of scientific terms. In this book the term "prior information" is often used. To statisticians this term means the information contained in n initial experiments as related to the $(n + 1)$th one. In a particular case $n = 0$, and then the term "prior information" will mean the knowledge gained by the experimenter before this series of experiments, from some quite different experience, related in some way or another to the problem under consideration. But imagine that in the audience where the lecture is being delivered by a statistician a philosopher is present. He will be irritated and decide that here the restoration of neo-Kantianism is being propagated. Indeed, the term "à priori statements" was introduced by Kant, who opposed it to the term "à posteriori statements." Kant needed this confrontation in order to develop the notion of inborn ideas. By now, the epistemological sense of these terms has been pushed somewhere to the background for all non-philosophers, but in reading certain texts it is restored to life immediately. If in our texts we deliberately underline the non-Kantian meaning of the term "à priori information," still this will not make it indubitably precise. In the literature on mathematical statistics there are many shades in the interpretation of this term. Sometimes attempts are made in the direction of their classification, but in vain. It is useless to introduce many narrow notions instead of one broad notion; it will only make our speech clumsy and complicated. The latter assertion is not only a phenomenon observed, but also a normative statement. It is impossible to imagine the existence of a multitude of narrow notions on a fuzzy field of meanings: they will be unavoidably hard to distinguish.

Two other terms widely used in modern logic are also connected with the name of Kant: they are "analytical statements" and "synthetic statements."[1] Now they are used not in an epistemological but only in a logical sense. One and the same phrase may be regarded at one time as an analytical statement and at another time as a synthetic one. And still some vague connection with Kantian notions remains associated with these terms. This connection is the inner succession of thought. And it is one of the functions of scientific terminology to preserve such succession in some concealed, unobtrusive form. Developing new concepts, we always confront them with the old, well-known structures and thus add certain inconsistency to the new statements.

The polymorphism of some new terms has been subjected to special study. The term "model"—quite a fashionable one nowadays—was honored by such a study in the paper by Chao Yuan-Ren (1962). He gives a list of 30 synonyms, i.e., of characteristics of "model," and of 9 non-synonyms, i.e., of notions contrasted to "model." We see that the synonyms of one and the same word are not always synonymous to one another, and sometimes a word is not even a synonym to itself. It is interesting to trace the historical development of the meaning of the word "model" (see, for example, Hornbey's dictionary). In English it means something ideal or perfect. In mathematics the word "model" was apparently introduced by F. Klein in the 1870s and later by Russell. One of the applications of this term in mathematics is connected with the concept of relative consistency. Above, it was already mentioned how a new system of axioms such as the axioms of Riemannian geometry is simulated on the spherical surface in the three-dimensional Euclidian space. Thus, Riemannian axioms turn into the theorems of Euclidian geometry, and hence it follows that the Riemannian postulates are consistent if the Euclidian geometry is also consistent. Further, the Euclidian postulates, according to Hilbert, are fulfilled on a certain algebraic model and, consequently, are consistent, if the same is true about the algebra. To this extent a model turns out to be a set of things, for which properties and relations are given by a certain theory—the theory which is being simulated. One and the same theory can be simulated on different ob-

[1] In the transcendental logic of Kant, such statements are called the synthetic prior statements, which remain prior, i.e., given outside experience, despite the fact that in such statements the predicate is not included in the subject; they are inborn statements. In modern logic, Carnap considers synthetic those statements which contain certain information about the external world: they are juxtaposed to analytical statements—tautologies—the truth or falsity of which does not depend on the connection with the external world. The division of statements into analytical and synthetic, despite its conditionality, proves very useful in the logical analysis of scientific texts.

jects. At present, we often ascribe a quite different meaning to the term "mathematical model building," meaning a certain simplified and rather approximate mathematical presentation of a complex system (Nalimov, 1971). The word "model" in this case is opposite to a law of nature describing phenomena in some rigorous way. One and the same system may be described by different models, each of them reflecting only one side of the system under study. If you like, this is the view of a complex system from a certain predetermined and apparently narrow angle. In this case, evidently, the problem of discrimination does not arise; different models may exist concurrently. To a certain extent, model, in this sense, behaves in the same way as the system it describes; yet the model is not identical to the system described. In linguistic terminology it must be said that a *mathematical model* is no more than a *metaphor*. This interesting idea was suggested by Hutten (1956). Now the question can be asked: Why was it only recently that building mathematical models of complex systems (such as those we come across in industry or in biology and sociology) became possible? No essential, hitherto unknown mathematical ideas appeared. The answer is simple: the psychology of research workers had changed. The standards for mathematical descriptions of the external phenomena have become lower. From the status of law they changed to the status of a metaphor. And psychologically we are quite ready for the possibility of using metaphors in science. All the arguments as to the possibility of applying mathematical methods in sociology are brought into focus by use of the word "model." If it is understood as something similar to the laws of nature, then nothing can be accomplished; if it is understood as a metaphor, than all the objections are eliminated at once.

Let us return to the confrontation of the two basic approaches to the notion of model in mathematics. In mathematical logic, the word "model" means the interpretation on a certain set of objects. One and the same theory may, as mentioned above, be simulated on different objects. Here we observe a multitude of interpretations, but this multiplicity is not of a metaphorical character. In applied mathematics, the word "model" means some theory of a complex system expressed in mathematical language. In this case one system is simulated by different models – by theories, and these models behave like metaphors. As we see, the difference in the understanding of terms proves very profound.

Sometimes in one and the same field of knowledge, and even in the same texts, we have to use the term "model" despite meaning by it quite different things.

Such contradictory reading of the term "model" in mathematical linguistics also has proved interesting. Trying to introduce strict and faultless clarity in his reasoning, Shreider (1971) supplied this term with three indices *m, l,* and *c*. In his interpretations, the term "model_m" corresponds to the strictly defined notion of model in mathematics; roughly speaking, this is an interpretation of the theory. The term "model_l" is the notion of model in linguistics, that is, the theory itself, or some hypothetical scientific construction. It turned out that the relation "to be model_l" is inverse to the relation "to be model_m." And lastly, the term "model_c" is a cybernetic understanding of this word. It had been proven that

$$\text{model}_c = \text{model}_m \text{ for } \text{model}_l$$

i.e., "the model_c of a real object is a mathematical model ("model_m") of a theory ("model_l") of this object." The trouble is that if we ascribe individual indices to all other possible meanings of the sense of the word "model," then no doubt nobody will be able to use the word. In speech (especially in reports and lectures), we always use the word "model" in various meanings and interlocutors understand this. But they will hardly understand anything if we speak as follows: "Having built a model in the fifteenth sense of the word, we have achieved the understanding of the word "model" in its twenty-seventh meaning." And still, an analysis of word meaning similar to that carried out by Shreider often proves very useful, since it allows us to penetrate more deeply into the polymorphous meaning of the word though it does not allow us to cope with it.

In the appendix to this book, a list of definitions of the term "statistics" is given. First of all, it is interesting as an illustration of the enrichment and broadening of word meanings. This word first appeared in fiction (Yule and Kendall, 1950): in *Hamlet* (1602, act 5, scene 2), in *Cymbeline* (1610, 1611, act 2, scene 4), and in *Paradise Regained* (1710, book 4), but the meaning of the word is not quite clear. It seems to be derived from the Latin "status" which means "political state." Later, the term "statistics" appears in science as well. Roughly speaking, three basic stages may be traced in the evolution of its meaning. First, it was the collation of data about the economic condition of a country based upon the analysis of those economic factors which can be expressed quantitatively. Perhaps in this meaning the term became connected with the German word *Staat* or the French word *état*, both of which mean "state." In the second stage of development, the term "statistics" was used for denoting the processing of any quantitatively expressible data, no matter the source: in science or in any other field of human activity. At this stage, statisticians were not worried about the reason and the way the data had been obtained. Nowadays, the term "statistics" is sometimes defined very broadly—as a metascience. The object of this science is logic and methodology of the other sciences, the logic of decision making in other sciences, and the logic of experiment. But such a broad interpretation is by no means widespread. At present we can still hear that the methods of statistics should be used cautiously, keeping in mind the priority of the quantitative over the qualitative. If some statisticians consider it senseless to divide statistics into mathematical statistics and statistics as a social science, others think such a division obviously necessary. Sometimes the aim of statistics is stated to be decision making under conditions of uncertainty. In a way, this definition is narrower than the definition of statistics as a metascience: it does not take into consideration all the questions connected with the logic of the sciences which are the concern of a metatheory. But at the same time it is broader, for it embraces both the problems of game theory and the problems of decision making in business. In this connection I should like to emphasize that the argument about the meaning of the term "mathematical statistics" is not a mere discussion about the limits of this or that scientific discipline. It is something much wider: it is the consideration of one of the problems of the philosophy of science. The discussion about the role of a probabilistic approach in grounding the methodology of scientific research has turned out to be an argument about the meaning of the term "statistics." In this respect it is especially interesting to observe the sharp divergence in estimating the role of the large-numbers law in social phenomena made clear in the articles by O. Yakhont and F. Lifshitz. This is not only the difference of opinions of the two authors but something much more significant since these opinions are given in the two leading Soviet encyclopedias: *Philosophical Encyclopaedia* and *Large Soviet Encyclopaedia.* It is noteworthy that the corresponding volumes of the two were published in the same year.

A collection of the definitions of the term "information," one of the main notions of cybernetics, would be of the same interest. A sampling of such definitions follows.

> Information is a name for the content of what is exchanged with the outer world as we adjust to it, and make our adjustment felt upon it. (Wiener, 1954)
>
> Information is . . . an attribute of objects, phenomena, processes of objective reality, man-made control computers, which consists in the ability of perceiving the internal state and the influence of the environment and preserving its results within a certain period of time; the ability of transferring the knowledge of the internal state and the obtained data to other things, phenomena and processes. (Kondakov, 1971)
>
> Information is the objective content of the connection between interrelated material objects, which manifests itself in the change of the state of these objects. (Mikhailov et al., 1968)
>
> Information is a philosophical category, considered side by side with

> such notions as space, time and matter. Most generally information may be represented as communication, i.e. the form of a condition between the transmitter sending the message and the recipient perceiving it. (Vorobiev, 1971)
>
> Information. The knowledge (in Russian "svedeniya") contained in a given speech excerpt and regarded as the object of transmission, storage and processing. (Akhmanova, 1966)
>
> Information means order; to communicate means to create order out of disorder or at least to increase the degree of order that existed previous to the message received. (Hutten, 1967)

Even this small collection of the definitions of the term "information" demonstrates how polymorphic this word is in its range of meanings. Here, the development of polymorphism is primarily connected with the fact that none of the definitions corresponds to our intuitive understanding of the meaning of the word. And any attempt at defining ascribes some new features to this word, features which do not clarify but, on the contrary, make narrow and thus obscure its sense, and indubitably increase the word's polymorphism. For example, Mikhailov's definition connects this term with material objects in the most rigid way and thus excludes from the term "information" our idea of theorems in mathematics, in proving which the material objects by no means interact. The notion of information as objective content of the connection between interacting material objects makes us exclude music from this category as well, for it is hardly of an objective character. The desire to regard information as a philosophical category, similar to space and time, throws us back to Kant's era. At any rate, now physicists are not prone to consider space and time as philosophical categories. In Akhmanova's definition the international word "information" is replaced by the Russian word "svedenija," the meaning of which is not further explained.[2] Strict limitations are imposed upon the word "svedenija": not all the "svedenija" appear to be information, but only those which are contained in a given speech excerpt. Non-speech excerpts, e.g., the results of observations presented as curves or in a discrete code on a magnetic tape, turn out to be excluded from the concept of information. Hutten's definition sounds the most pleasant. It does not encompass the depth of the notions connected with the term, but it does reflect the content ascribed to it by physicists and, I dare say, experts in cybernetics. It is noteworthy that this definition sounds similar to the oldest idea of the role of a word in the creation of the universe. In the Gospel according to St. John we read: "In the beginning was the Word. . . . All things became through Him; and

[2] It is not easy to translate the word "svedenija." I have translated it above as "knowledge," but this translation is not quite adequate. It is better to translate this word as "information," but in this case Akhmanova's definition becomes a tautology.

without it did not anything become: that which became . . ." In the modern canonical versions of the Gospel of St. John the very polymorphous Greek concept λογος is traditionally translated as "word," and then the word acquires the role of a constructive and arranging force. It is in this sense that the word "information" seems to be understood nowadays.

It is interesting to call the reader's attention to the interpretation of the term in the *Philosophical Encyclopaedia* (*Filosofskaya Entsiklopediya*, 1962). It runs as follows: "Information (Latin—informatia)—see: information theory." Further, under the heading "Information theory" the questions with which this theory deals are enumerated, and various presentations of the quantitative estimation of information as a measure of order are given: according to Hartley and Shannon; to R. Fisher and A. N. Kolmogorov; to N. Rashevsky; to R. Carnap and Y. Bar-Hillel; and, at last, to Yu. A. Shreider. The definition of the term "information" proper is not given at all. The picture becomes very curious: a non-philosopher (Vorobiev, 1971) puts the notion of "information" under the heading of philosophy, but philosophers refuse to consider it from the philosophical point of view.

The above examples seem sufficient to back up the correctness of my thesis about the deep polymorphism of the language of science.

In the language of science, polymorphism manifests itself more clearly than in ordinary language. The reason is that here the words encode whole concepts. Scientific concepts may be very fuzzy and versatile. Many scientists understand only certain aspects of the complex system of notions. The above definitions of the word "information" are just a collection of judgments on different facets of the complex system which has recently been crystallized out in a separate scientific trend called cybernetics. The same is true of the term "mathematical statistics," and to a lesser degree of the terms "model" and "prior probability." The latter two terms encode not just one large but several small interrelated concepts. Such seemingly simple physical notions as mass and force are also theory laden. Here, I shall refer to Einstein and Infeld (1954): "Physics really began with the invention of mass, force and an inertial system. These concepts are all free inventions" (p. 295).

Concepts cannot be defined; they should be explained. The conceptual character of terms creates intensified polymorphism of the language of science. The deeper and more complicated the concepts encoded by the term are, the greater its polymorphism.

Here, a rhethorical question may be asked: If the polymorphism of language both in science and in ordinary speech grows with time, then won't language degenerate in the future, i.e., each notion will become all-embracing, and all the notions will have the same meaning? Indeed, asymptotically it may seem so in our model. But we have already agreed to regard language as a living organism, and like any organism, with age it must give way to another one. By the way, aging is no more than a natural process of information storage, which prevents further progress. The increase in the polymorphism of words does not go on smoothly. If the meaning of a word is imagined as a continuously widening field, then at certain moments part of this field may be lost, i.e., forgotten. Words undergo a complicated process of development; interesting examples of the semantic history of words are given in Budagov's book (1971).

Speaking about scientific terms, we must pay attention to another peculiarity. *In the process of the development of science, its words gain prestige.* Furthermore, the same occurs in social life, but here we shall restrict ourselves to the analysis of scientific terminology. When a scientist proposes a new concept, he wants to express it in old words. If he manages to do that, the new theory immediately gains the prestige already associated with these terms. If, for example, an absolutely new meaning is ascribed to mathematical statistics, it is considered a metascience, and this new meaning is put into the old word "statistics"—a word which has already gained very high prestige. Now imagine that, developing a new concept, a scientist expresses it in new words. It will be equivalent to the loss of the game. More conservative colleagues of his will declare: "He says something entirely different about the problems we are concerned with." In Russia the word "statistics" so far has been strictly connected with economics, and to avoid depressing arguments I have suggested calling this new understanding of statistics by a new term: "the mathematical theory of experiment." The arguments about words in science, which irritate many of us, are sometimes not at all small talk. (The ideas developed in this paragraph were suggested by S. K. Shaumyan during our discussion of the manuscript.)

Specific languages of science, their slang character. To some extent the languages of science are organized in a way similar to the thieves' cant. In both cases the words and grammar of everyday language are used; it is seldom that new specific terms are introduced. These new terms and the new meaning ascribed to old words borrowed from the everyday vocabulary give an esoteric character to slang language: they prove comprehensible only to the initiated. And still, the similarity between the language of science and slang is only superficial, so it is better to speak of *specific languages of science.*

My understanding of the specific character of the language of science can be illustrated by an example. After a report on mathematical statistics to an audience of engineers in metallurgy and the science of metals, one of the listeners said that all this was certainly very interesting but, unfortunately, incomprehensible. Rather irritatedly, he added: "And why not change such incomprehensible terms as regression, correlation, variance into simple Russian words?" The lecturer answered: "Then let us give up such well-known terms as martensite, troostite, crystobalite, and substitute for them such simple Russian words as ticks, crosses, dashes, or dots according to what is seen under a microscope when examining metallographic and petrographical sections." This suggestion irritated the audience, and the reason is clear: the point is not that "perlite" and "martensite" are foreign words. The difficulty is of quite another origin: these words encode complicated metallographic concepts, and if we give up this system of codes and turn to the arbitrary but apparently understandable words, then in conversation we shall have to explain all the concepts from the very beginning. In the same way in mathematical statistics, the terms "variance" and "regression" encode whole scientific concepts, and the lecturer's difficulty lies in explaining them popularly and using them for developing the ideas which he wishes to state in his lecture. If a reader meets an unknown term in a paper on mathematical statistics, an explanatory dictionary of specific terms would be of no use for him because it is not a strict definition of the term (if it does exist) that it is important to know but all the concepts connected with it. Thus, such a language barrier may be also called a conceptual barrier. In contrast to ordinary human language, the language of science is of a much more distinctly *coded* character. The depth of coding or, in other words, the informational capacity of terms grows in time with the development of scientific concepts. The difficulties are also redoubled by the fact that specific languages often use the words of ordinary language in a special sense. For instance, everybody knows the common meaning of the word "replica," which is a French borrowing. In mathematical statistics there are such terms as "replica," "fractional replica," and "regular replica" which have a specific meaning. "Fractional replica" means some specially selected part of the complete factor experiment—its fractional, i.e., partial, repetition. The meaning of this term becomes clear after substantial acquaintance with the concept of experimental design. Finally, in optics "replica" is a copy of a diffractional lattice prepared in a special way. All three of these terms with different meanings originated from the French "réplique."

The slang-like character of speech manifests itself not only in the sciences but in the humanities as well. The Russian edition of this book contains several extracts from the reports made at a conference on the Oriental problem. These examples deal with specialized expressions that cannot be translated into English. The paradox is that sometimes foreign words are inserted into Russian speech in an extraordinary manner which makes it sound elegant and artificial. These word combinations are actually words with Greek and Latin roots, and substituting for them words with Russian roots will eliminate the effect.

Recall once more that specific languages of science are continuously changing. New concepts emerge, and old notions are often assigned a new meaning. Because of its continuously changing nature, scientific language is accessible only to those working in the field and thus constantly interacting with the informational flows in science. The same phenomenon, but to a lesser degree, can be observed in ordinary language. Suppose you give a foreigner who has been living in the Soviet Union for a long time a magazine in his native language and ask him to translate several pages. He will translate ordinary text easily, but will immediately stumble over new slangish or idiomatic expressions, and cartoon captions will pose almost insuperable difficulty for him: as a rule they are based on certain peculiarities of current life, encoded in specific words. No matter how long I study English, I shall probably never learn it to such a degree as to be capable of translating a caption under a cartoon in such an intellectually respectable American magazine as *The New Yorker.*

Babelian Difficulties in Science

With the development of science, more and more separate specific languages of science have crystallized. This facilitates the exchange of information on the borders of narrow branches of science, but hampers mutual understanding between neighboring fields of knowledge. If I am permitted to be a bit frivolous, I shall say that the situation is the same as that at the building site of the Tower of Babel.

Heated discussions as to whether or not this or that field of knowledge can be considered an independent discipline are common occurrences. In discussing this question the opponents define various criteria. One of these is the statement to the effect that every independent scientific discipline should have a research method of its own. I think that, proceeding from the above, another quite simple criterion may be suggested: the emergence of a new independent scientific discipline must be accompanied by the emergence of a new specific language (or, rather, a dialect). The emergence of essentially new problems immediately leads to the emergence of the new language in which they are discussed. In contrast, the creation of a new science is not necessarily accompanied by the creation of new research methods, especially nowadays when many new branches of knowledge appear at the junction of previously existing ones and use their research methods. For example, the design of experiment – a subdivision of mathematical statistics – seems to me to be a new independent scientific discipline. In the process of its development, solving its own specific problems, this discipline has developed its own specific language. This language irritates specialists in mathematical statistics if they are not specialists in the problems of experimental design as well. At the same time, this new discipline has no unique methods. It uses methods commonly used in mathematics: linear algebra, combinatorial analysis, numerical methods of analysis, and, in its most unique manifestations, the methods of functional analysis, set theory, and abstract algebra. The above statements sound similar to those of Shreider (1969), who asserted that we should ascribe the greatest profundity to the truths which change the human thesaurus to the greatest degree.[3]

We often hear discourses on differentiation and integration of science. The process of differentiation can be easily traced by the emergence of new local, specific languages of science. As far as the integration of science is concerned, it is wishful thinking rather than an actually observed phenomenon. If this process had taken place, then we should have noticed at least some vague signs of the emergence of a language necessary for it. How can we speak about the existence of a specific manifestation of scientific thinking if there is no language in which it can be expressed briefly and clearly? The only phenomenon we can observe now is the appearance of new branches of knowledge at the intersection of some already existing disciplines which seem to have nothing in common. This is not integration but additional differentiation of knowledge. Every newly created discipline of such a kind is clothed in the attire of a newly created language. Here again, we refer to the example of the emergence of a new branch of knowledge called "experimental design." This branch has appeared at the junction of many subdivisions of mathematics, but it has not led to their integration. Metamathematics – a science dealing with the foundations of mathematics – cannot be regarded as a discipline resulting from the integration of mathematical knowledge. It is just a new subdivision of mathematics with its complicated concepts and its own specific language in which its concepts are encoded. This subject is remote from the representatives of other subdivisions of mathematics. Similarly, modern logic cannot be regarded as a result of the integration of differ-

[3] This is a one-way criterion: if in any field of knowledge there appears a new specific language, it undoubtedly means the appearance of a new scientific discipline, but scientifically formed languages may emerge with the construction of a system of notions in a region which is far from being scientific. Freud's theory is an example; it has a scientifically formed language of its own, but, strictly speaking, it is not scientific for it is formulated in such a way that it cannot be verified. This is not to say that I have a negative attitude to this theory; besides, I do not think that human intellectual activity should be completely reduced to scientific categories. A characteristic feature of science is the possibility of verification of its hypothesis (although, strictly speaking, it is difficult to give a clear-cut definition of what we understand by the term "the possibility of verification").

ent branches of knowledge. Its language is as specific as the languages of other branches of knowledge, and quite a large group of scientists fail to understand it.

Specific languages of science have another function. Delicate refinement of language turns out to be a form of scientific aristocratism, a sign of belonging to a certain scientific community similar to the situation in old Russia when speaking good French indicated that one belonged to the nobility. The representatives of some fields of knowledge, especially mathematicians, or at least some of them, have always considered themselves to be at the Olympus of science. A young mathematician thinks that by vulgarizing his language he betrays the refinement he has been taught and, consequently, loses the right to belong to the scientific community which it had been so difficult for him to enter (Nalimov and Mul'chenko, 1972). Strange as it seems, this aristocratism is also taught in our universities—God knows how. Unfortunately, it commonly happens that the superior verbal behavior of young mathematicians insults the representatives of other branches of knowledge who have come to them for consultation.

The Problem of Standardization of Scientific Terminology

I do not want the reader to get the impression that I reject the necessity of making scientific terminology stricter. From my concept that the polymorphism of language makes it a truly powerful means of communication, it does not follow that in scientific language we should permit that innumerable variety of terms which can often be observed.

Preobrazhenskaya et al. (1974) present interesting data which deal with the frequency of statistical terms in publications on spectrochemical analysis and analytical chemistry. One of the histograms from this paper is given in Fig. 5. Such graphs are interesting in two respects. Firstly, they permit us to judge the degree of penetration of statistical terms into this or that branch of knowledge. Here we see that serious concepts of mathematical statistics, connected with such terms as "regression analysis," "the least square method," and "distribution," are rarely used in the field under study. Furthermore, from this graph we see that to denote a single concept of "error" a variety of synonymous terms are used: ошибка (error), точность (precision), отклонение (deviation), погрешность (fault), воспроизводимость (reproducibility), расхождение (divergence). The picture will become still more confused if we consider the word combinations: точность анализа (precision of analysis), погрешность анализа (slip in the analysis), ошибка воспроизводимости

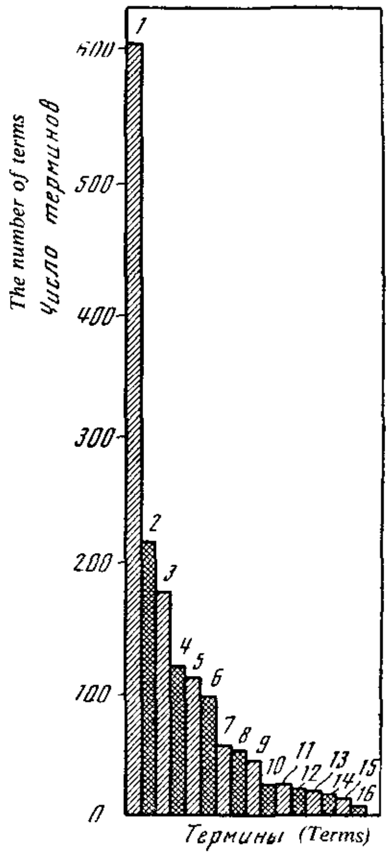

Fig. 5. *Distribution of terms of mathematical statistics according to the frequency of their occurrence in publications on spectrochemical analysis in the journal Industrial Laboratory (Slavnyi, 1969). (1) Error; (2) precision; (3) faultness; (4) deviation; (5) reproducibility; (6) variance; (7) confidence interval; (8) variance coefficient; (9) divergence; (10) criteria; (11) dispersion; (12) distribution; (13) correlation analysis; (14) least square method; (15) methods of mathematical statistics; (16) regression analysis.*

анализа (reproducibility error), отклонение результатов анализа (deviation of results), достоверность анализа (validity), оценка погрешности (error estimation), оценка достоверности (validity estimation), статистическая достоверность анализа (statistical validity), относительное расхождение анализа (relative deviation), средняя относительная ошибка (average relative error), относительная ошибка (relative error), случайная ошибка (random error), статистическая ошибка (statistical error), отклонение от истинного результата (deviation from the true value), относительное стандартное отклонение (relative standard deviation), относительная точность (relative precision), относительная величина дисперсии (relative value of variance — the expression is absolute nonsense), абсолютное расхождение (absolute deviation), средняя статистическая ошибка (average statistical error), относительная статистическая ошибка каждого измерения (relative statistical error of each measurement), вероятная *случайная ошибка* (*probabilistic random* error).

All these terms, formed by combining two or three words, are synonymous in the sense of "estimation of the degree of uncertainty, connected with the result of analysis." Naturally, the authors of all the publications investigated did their utmost to present their results in a form comparable to that used by other authors. But have they really reached this aim by using terms of such mixed and unintelligible character?

Recently, the State All-Union Standard No. 16263-70 was published in the Soviet Union with the title "State provisional system for the unity of measurements. Metrology. Terms and definitions." In this publication the term погрешность (mistakenness) is suggested as a standard.

Figure 5 suggests that it will hardly become established. This term, at least with specialists in substance analysis, is six times more rarely used than the unrecommended term ошибка измерения ("an error of measurement"). It seems useful to dwell upon the analysis of this Standard in more detail. In the Soviet Union it is one of the first attempts to *decree* the language of science in that branch which deals with an activity of all experimenters, and such a standard has the status of law. The preface to the Standard reads as follows: "The terms established by the present Standard, are recommended for use in documentation of all kinds of manuals, teaching aids, technical and reference literature." This Standard was developed by serious scientific institutions: Mendeleev All-Union Scientific Research Institute for Metrology and All-Union Scientific-Research Institute for Technical Information, Classification and Codification. Still, we find quite strange recommendations in it. One such oddity is the confrontation of the terms наблюдение (observation) and измерение (measurement). It runs as follows:

> Результат наблюдения (The result of observation). The value obtained by a single observation.
>
> Результат измерения (The result of measurement). The value obtained by means of measuring it.

Further, two more notions are introduced: среднее квадратичное отклонение результата наблюдения (the average square deviation of the result of observation) and среднее квадратичное отклонение результата измерения (the average square deviation of the result of measurements). One can hardly understand when and which of these terms should be applied.

It is also strange that in this Standard the terms *точность измерения* (precision of measurement) and правильность измерения (accuracy of measurement) are confronted. The following definitions are given there:

> Точность измерения (Precision of measurement). Quality of

> measurements, reflecting closeness of their results to the true meaning of the value measured.
>
> Notes:
>
> 1. High precision of measurements corresponds to small errors of all kinds, both systematic and random.
> 2. Quantitatively the precision may be expressed by the inverse value of the modulus of the relative error.
>
> Правильность измерения (Accuracy of measurement). Quality of measurements reflecting closeness to zero of systematic errors in their results.

Here everything is puzzling. How can two separate terms coexist if one of them, правильность (accuracy), is given by reference to the other one, точность (precision). In Anglo-American literature, the terms "precise" and "accurate" are traditionally contrasted: random error is connected with the first one, and systematic error, with the second. The two terms "random error" and "systematic error" logically pertain to notions of different types (in the sense of Russell), and to build here a combined notion is as strange as to say, "I see two objects: a chair and furniture." What seems even more strange is the statement about measurements having qualities which are defined by quantitative characteristics. The concept of the true value of the quantity measured is defined in none of the above definitions; it remains a vague, philosophically shaped notion. It is very surprising that both the concepts of точность (precision) and правильность (accuracy) pertain only to measurements but by no means to observation, though it follows from the same Standard that the observations are also expressed quantitatively.

I have dwelt in such a detail on this example of the terminological Standard to demonstrate how great the difficulties are which are faced in the attempt to make scientific terminology stricter. The above-mentioned Standard will hardly be of any use for Soviet science. Nevertheless, this is a curious precedent: scientists are officially presented with terms which are beneath criticism from the standpoint of logical analysis and which do not correspond to the historical traditions of the scientific community. I wonder what will come of it.

We can formulate the following sufficiently general statement: the broader a scientific term is, the more difficult it is to define it. Here is an example. In the Soviet *Philosophical Encyclopaedia* (the article "Experiment" by B. Dynin) the term "experiment" is defined as follows:

> *Experiment*—sensual-objective activity in science performed by theoretically cognized means.

Imagine that an experiment is performed which is aimed at the registration of infrared rays. The results of the experiment go into the computer, and the output order appears in mathematical language. What is to be considered "sensual-objective activity" here? Is it the obtaining of an infrared spectrum which we do not perceive through our senses? Is it correct to assume that the spectrograph, spectrum generator, and the aggregate for the registration of the spectrum are "theoretically cognized means"? The modern mathematical theory of experiment actually proceeds from the opposite assumption. It states that the experiment is being carried out in a situation which cannot be exhaustively described by theory. On this basis, it is suggested that the conditions of the experiment be randomized in order to avoid possible systematic errors. Randomization would be unnecessary if the experiment were performed in a situation absolutely under the experimenter's control. If the above definition is to be retained, then the largest part of scientific experimentation must be considered as non-scientific activity!

The difficulties in ordering scientific terminology seem enormous. Scientists want something to be done in the field, but such activity must be extremely cautious. To my mind, the terms should be explained rather than strictly defined. As a rule, every concept in science is closely linked with a field of meanings which has been formed in the course of a long history of development. Any attempt at rigorous definition may impose unwanted restrictions on the field. Following Spinoza, we may say that any definition is a negation—in our case, the negation of that part of the semantic field which has not entered the definition. Should such great restrictions really be imposed upon the semantic fields as has been done in the above example with the term "experiment" or still earlier with the term "information"?

Terms in science must serve not only for the expression of previously developed concepts, but also for the formulation of statements in the future. That is why scientific terms *must be open*. Even in mathematics, as was clearly demonstrated by one of the examples of Lakatos (1963–64), the criticism of the previously stated solutions leads to the broadening of the meaning of the conceptions.

In any case, it is clear that any terminologic recommendations must be preceded by substantial logico-linguistic analysis of the whole variety of actual scientific terms.

Contents

Chapter 4