

V. N. BELOOZEROV - A. A. POLIKARPOV

A MECHANISM OF THE ANALYTICITY DEVELOPMENT IN A
LANGUAGE. ELEMENTS OF THE THEORY AND THE
COMPUTER EXPERIMENT

1. From the point of view of the general cybernetic theory we consider language to be a device, which regulates the information streams in social systems.¹ On this basis it is possible to come to the conclusion of bilateral dependency of the language organization level (the level of its complexity) on (1) complexity value of the language mechanism input information stream, and (2) complexity value of the output information stream which is to be produced by the language mechanism.²

The relation between mentioned values may be represented by the following formula:

$$L = \sqrt[k]{S},$$

where L is complexity (volume) of the language (in the elementary language units);³ S is complexity (volume) of the input information stream (in elementary sense units, which are supposed to be encoded); k is complexity (volume) of the output information stream (in elementary sign units, in which sense units are coded).

In the given paper we consider certain moments of dependency of general level of human language complexity on some values of the

¹ The analysis of systems, in which communication processes are carried out and in which the necessity in communicative regulators arises, is made in our papers. It should be noted that our analysis is based on the ontological model of the communicative process, which is represented in the series of papers on systems linguistics by G. P. Melnikov, published in 1967-1972 (G. P. MEL'NIKOV, 1967, 1969, 1970, 1972 a).

² It follows from the fundamental cybernetic law of the necessary variety (W. R. ASHBY, 1958), which determines the functioning of any regulator. A more detailed analysis of the given dependency is made in the paper of A. A. POLIKARPOV (in press).

³ Every language unit consists of a model of a sign coupled with its meaning. On the principles of the inner organisation of language one may consult the papers of G. P. MEL'NIKOV (1969, 1970).

output information stream; internal typological reconstructions of language, wherein the change of its complexity is expressed, and some fundamental reasons of the changing in the values of the output information stream optimum redundancy are also treated.

As is going to be shown below, the changing in the values of that parameter depends on some specific processes in the social systems of communication. That is why the main problem of the given paper refers to the sphere of socio-linguistic problematics and constitutes a fragment of the more complex socio-linguistic model, which is being worked out at present.

2. Factors that lead to the changes in the parameter of language complexity are, in our opinion, connected with increasing or decreasing of specific "noise level" in language communication channels. Such "noise" appears, for instance, in the case of long-time and intense language contacts, when the community includes different language bearing communicants amalgamated. If the income of foreign-language communicants is weak and occurs from time to time, the language contact does not effect considerably the communication situation (the foreign incomers accept gradually the common (original) language). But in the case of stable income of such communicants into the communicative system, and for that reason of stable low average value of language message comprehension in the community, the communicative system has to change the values of its parameters in order to increase the efficiency of communication. One of the results of such a process is the reorganisation of language, its transition to the more simple organisation, which is described in the terms of linguistic typology as analyticisation.

The interrelation between the phenomenon of language contacts and language analyticisation was already indicated more than once (L. TESNIERE, 1939; U. WEINREICH, 1953; U. WEINREICH, 1972; G. ŠUCHARDT, 1950, V. JU. ROSENZWEIG, 1972) but the mechanism and transitional stages of the development of the process remains undiscovered up to the present time. The communicative and intralanguage mechanisms of analyticity developing in language become to our mind, apparent, if we consider the following hypothetical model.

3. We take as the starting point of building the model the fact that the average level of mutual comprehension of the interlocutors decreases at the beginning of interlanguage contacts. Due to the vital

importance of achieving mutual comprehension the interlocutors naturally try to make their speech explicit to each other. This is achieved by the increase of volume, complexity of messages, which as regards the preceding form of organisation of sign messages, can be defined as the increase of message redundancy. After a certain period of time the repeated attempts of the interlocutors to establish communication result in discovering by the communicants a certain part of the common language vocabulary as the most accessible for the typical interlocutor under given circumstances, i.e. they result in the narrowing of the vocabulary in comparison with the vocabulary of the language at the previous stage. It turns out that the most suitable and productive part of the vocabulary is, as a rule, that part in which the most general, most universal, most frequently used words are concentrated. This very part of the vocabulary allows, in a general way, to "hint" each phenomenon of reality. Besides, mastering of the language system begins only with this most frequently used, most universal part, even at the lowest stage of one's learning the language.

As a result of the universalization of a definite part of meanings or escaping of some of them, all the system of semantic oppositions in the language is being reconstructed, i.e. the most peripheral oppositions fall out and the central ones are strengthened. This process in the language is well exemplified in the escaping semantic oppositions between different grammatical "words" (grammatical morphemes and pronoun-adverbial indicators). As a result of this, many grammatical indicators stop "working" and, being useless, are reduced, and words shortened till one syllable only. The stabilising level of an optimum redundancy of the sign messages in these conditions, beyond doubt, is higher than during the period, which preceded the reconstructions of the language system, because in order to reach the optimum (i.e. preceding) level of understanding with the help of the most general in their meaning sign-"hints", it is necessary to describe the object of a message from several various points. Aims, which are reached by a synthetic language with the help of a single but complex "hint", in analytical languages are reached with the help of several, actually grouped, "hints", which are not necessarily fixed beforehand in the speaker's memory to be used together. Meanwhile, the synthetic language organization presumes that the "hints" are fixed in the long-time memory of a typical narrator. In the latter case, the language intensively spends the memory volume and spares the operating efforts for message producing. On the contrary, the analytical language forces

to economize the communicants' long-time memory and to exploit their efforts when encoding the communicative intentions, to a greater extent.

Thus we may assume that the synthetic type of language organization is preferable because of communicative causes. This preference arises from the fact that the circumstances accompanying this language type allow the fixing in the language system of many of the most successful means of denoting objects of reality. Analytical type of language organisation, on the contrary, makes the average language speaker fix in the language only a certain minimum of language experiences, it makes him improvise the actual formation of the message. This, apparently, does not promote the precision, optimality in denoting objects.

4. We undertook a mathematical check-up of the connection between the starting point of the process of analyticity increase (i.e. the fact of decrease of the level (probability) of comprehension of the sign messages generated with the help of the language) and one of the terminal moments of the process (the increase of the optimum redundancy level of messages built on the basis of the preserved part of the vocabulary).

For the calculation, the following model of communication was designed. A speaker must "convey" to a listener the sense, which can be "transferred" in the most favourable, so-called "noiseless" case, with the help of the utterance containing N signs. The average probability of comprehending the signs (with definite "interferences" present) is P . From the whole number of signs, which to some extent make a "hint" to the sense, the speaker chooses m items ($m \geq N$), and then he produces the utterance, consisting of the signs chosen before, to the listener. Judging by the reaction of the listener, he makes a conclusion on the result of the experiment, whether the sense was caught or not. The sense of the utterance is considered to be understood if the number of signs, clear for the listener, exceeds N . Then the act of communication appears to be accomplished. If the sense of the utterance has not been caught, the speaker repeats the random choice of m signs from the same set, and build new utterances until the listener comes to understand the sense of the utterance at last. In this process the previous utterances are considered to have neither effect on the comprehension of further "transforms" nor on the new choice of the signs.

In such a model of conversation, the time necessary for one act of

communication can be evaluated by the expected number of signs (K), produced in all the utterances taken together. If the average probability of every sign to be understood is P , then the probability of comprehension of at least N signs of the whole number (m) is $P_{mN} = \sum_{i=N}^m C_m P^i (1-P)^{m-i}$.

The average number of utterances, given in order to reach comprehension, is

$$\sum_{i=1}^{\infty} i P_{mN} (1 - P_{mN})^{i-1} = 1/P_{mN}$$

Then the expected duration of the communication is

$$K = m/P_{mN}$$

This value was calculated with the help of the computer for different magnitudes of m , N , P . The typical diagrams of dependency $K - m$ for different P is represented in Fig. 1.

The optimum utterance redundancy minimizing the expected time of reaching comprehension appears to exist at any volume of information, which is to be transferred, and at any probability of comprehension of the signs transferred. The diagram showing the dependency of the optimum redundancy $A_{opt} = \min \frac{m-N}{N}$ on P is represented in Fig. 2.

The data given above show that, for the communicational situation characterized by the partial comprehension of signs, the more redundancy the message contains the more is the probability of miscomprehension. The speaker's intuitional search for the optimum strategy of communication, leading to the rise of sign saturation of utterances, evokes other phenomena, characterising the development of analytic features in language; the higher the degree of these, the lower is the level of comprehension.

To sum up, we must emphasise, that on the basis of the model given above a more complicated model might be constructed. The phenomenon of communicants' adaptation to each other in the course of the attempts to "transfer" the sense information must be taken into consideration in further research.

The further prospects of elaboration and computerized checking of the model seem to be the following: the comprehension level during the communication can be connected with such parameters

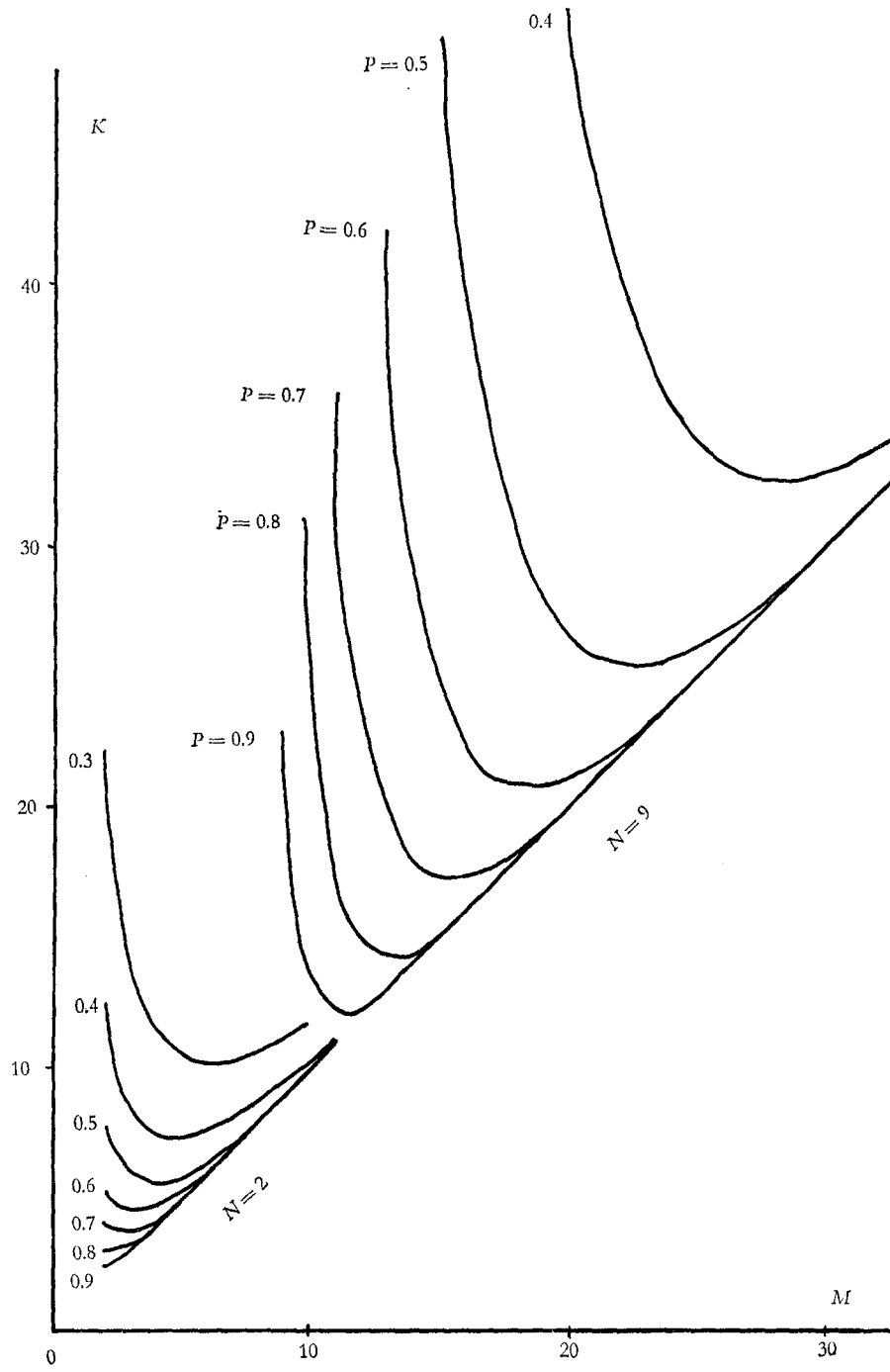


Fig. 1.

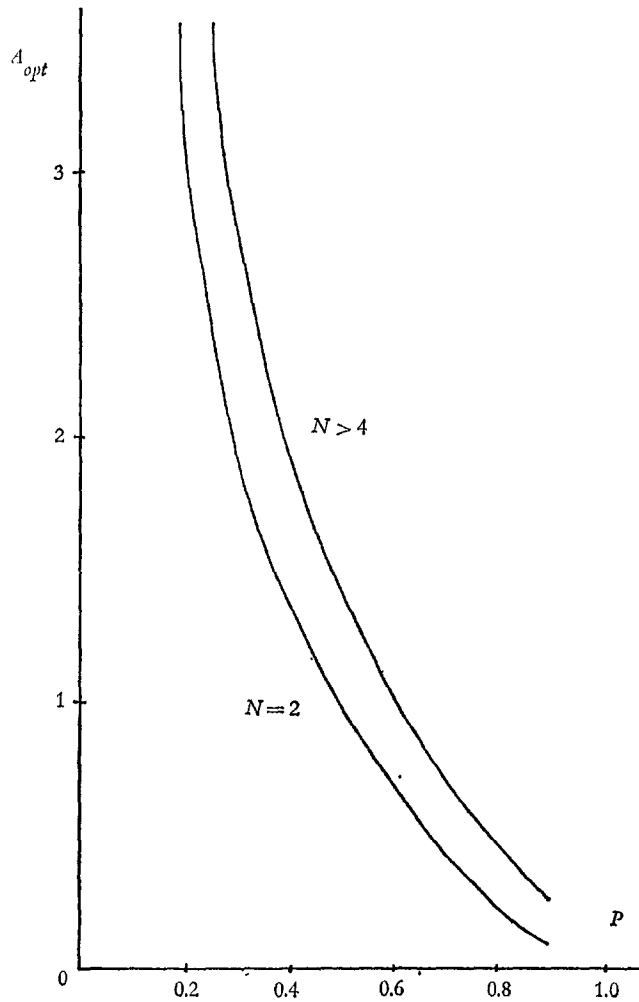


Fig. 2.

as (1) the degree of divergency between the language units stock of contacting languages (the more prominent is the divergency between contacting language the lower is the level of comprehension), (2) the percentage of nonprinciple language speakers within the community (the more numerous the representatives of the language, the lower the average level of comprehension), (3) the intensity of contacts between the representatives of different languages (the more frequent the con-

tacts between different languages representatives, the lower the average level of comprehension in that communicative system), etc. Further, the definite qualitative types of languages⁴ can apparently be evaluated by means of the definite quantitative index.⁵ The efficient shift of the quantitative index during language evolution, breaking the definite ranges boundaries, seems to define the transition of language into another qualitative type. The quantitative characteristics of languages can be calculated theoretically and empirically checked up by numeral characterisation of communicative situations (percentage of nonprinciple language representatives, the divergency degree of contacting languages, etc.) in which the language is formed. As a result, the matrix of language types can be associated with the matrix of communicative situation types.

To put it briefly, the given perspectives are those of creating a complex-structured sociolinguistic model with many parameters reflecting the data of modern qualitative and quantitative linguistic typology as well as quantitative and qualitative sociological typology. One may expect that the successive detection, the quantification of the values of linguistic and sociological parameters, the revealing of their relationships will result in building up the explanatory sociolinguistics theory. This theory can find out the correlation between the given states of a language and the given states of society employing it. This theory is very likely to predict with a high degree of certainty the future development of languages and the human language in general with a view of perspectives of humanity's social development. It is possible, having a carefully worked out theory (that is, knowing principle correlations between language and society), to simulate with the help of a computer the evolution of the given language as the function of the facts given by the history of the people employing this language, of the communicative system history, in which the language is formed.

⁴ Turn to one of the most full, in our opinion, classifications of language types, represented in the paper of G. P. MEL'NIKOV (1972 b).

⁵ In the first line we mean the criteria of evaluation of real object complexity based on the categorical system of information theory. These problems are being considered, for example, in the paper of S. M. DANCOFF and H. QUASTLER (1953).

REFERENCES

- W. R. ASHBY, *An introduction to cybernetics*, London 3rd imp., 1958.
- V. N. BELOOZEROV, *O fonologičeskich edinicach v kanalach massovoj komunikacii*, in *Semiotika sredstv massovoj komunikacii*, Materialy simpoziuma 18-20 ijunja 1973 g., MGU.
- S. M. DANCOFF, H. QUASTLER, *The Information Content and Error Rate of Living Things*, in *Information Theory in Biology*, Urbana 1953.
- G. P. MEL'NIKOV, *Sistemnaja lingvistika i ee otnošenie k strukturnoj*, in *Problemy jazykoznanija*, Moskva 1967.
- G. P. MEL'NIKOV, *Sistemnyj analiz čelovečeskich jazykov v svjazi s problemoj modelirovanija jazykovych processov i postroenija universal'nych mašinnych jazykov*, Otčety LASURS MGPI im. V. I. LENINA, Sektor obščej teorij sistem, za 1968 i 1969 gg., Moskva.
- G. P. MEL'NIKOV, *Rečevaja dejatel'nost' pri bilingvizme*, in *IV seminar po voprosam teorij i praktiki perevoda naučnoj i tehničeskoj literatury*, 2-4 nojabra 1970, Tezisy dokladov, Moskva-Tbilisi 1970.
- G. P. MEL'NIKOV, *Sistemnyj podčhod v lingvistike*, in «Sistemnye issledovanija», Ežegodnik 1972, Moskva 1972(a).
- G. P. MEL'NIKOV, *Determinantnaja klassifikacija jazykov i jazyki bantu*, in «Africana. Africanskij etnografičeskij sbornik», IX, Leningrad 1972 (b).
- A. A. POLIKARPOV, *K objasneniju osobennostej sintaksisa jazyka nauki kak jazyka massovoj komunikacii*, in *Semiotičeskie problemy jazykov nauki, terminologii i informatiki*, Materialy simpoziuma. t. 2, MGU, 1971.
- A. A. POLIKARPOV, *O meste sistemy massovoj komunikacii [MK] v tipologii komunikativnych sistem*, in *Semiotika sredstv massovoj komunikacii*, Materialy simpoziuma 18-20 ijunja 1973 g., MGU.
- V. JU. ROSENZWEIH, *Jazykovye kontakty*, Leningrad 1972.
- G. ŠUCHARDT, *Izbrannye stat'ji po jazykoznaniju*, Moskva 1950.
- L. TESNIERE, *Phonologie et melange des langues*, in «Travaux du Cercle linguistique de Prague», VIII (1939).
- U. WEINREICH, *Languages in Contact*, New York 1953.
- U. WEINREICH, *Odnodzycie i mnogojazyčie*, in «Novoe v lingvistike», vyp. VI, Moskva 1972.

