Towards a Semantic Theory of Information

Ricardo Pérez-Amat García (posthumous)

Being professor at the University Juan Carlos I (Madrid, Spain), he passed away in November the 14th 2008. This work was posthumously translated by Prof. B. Al Hadithi and J.M. Díaz Nafría.

Abstract: Information can be understood as that which reduces uncertainty, no matter what origin it has. In the field of human communication, information is only meaningful if it is part of a finished or intentional action. Meaning should be gathered from the empirical perspective of the use of language.

If we study the processing of signification through transmission of the normal use of language, we will see that it takes place communicating a set of prototype categories, the core or central facts, which defines meaning as empirical hypothesis. But if there are central facts showing the use of words, then other facts –more or less peripheral– should also arise, whose knowledge is necessary in order to communicate in contexts far away from the "denotative conceptual norm". Hence meaning can be represented by a fuzzy subset of the universe of discourse partition set. This concept of meaning may be integrated in a formal model of semantic source and information may be measured by non-probabilistic entropy.

Keywords: Semantic, pragmatic, fuzzy logic.

"Mathematics does not practically appear in physiology, ethology, psychology or social science, if it is not in the form of statistics recipes whose legitimacy is suspected. There is only an exception: the mathematical economy, with Walras-Pareto's model of economic change, leading to outline interesting theoretical problems (…).

(…)

Indeed, specialists know about this relatively quick degradation of the possibilities of mathematical tools (…), but the issue is rarely known by general public.

(…)

From the perspective of the internal use, (…) the techniques of approximated mathematics (approximation) make possible the emerging of a significant scientific production. Every attempt of quantitative modelling, either if it has foundations or if it is little or poorly founded, may encourage scientific publication. (Thom, 1984, p.139)

By starting this presentation with a reference to the mathematician René Thom, I would like to show from the very beginning –if you allow me to speak in first person– the theoretical and formalizing character of this investigation, yet being conscious of the varied risks –even practical– that such endeavour brings. If –as Paul Feyerabend (1981) asserted– the scientific activity has really something to do with a kind of life, there is probably no other way to understand my scientific and academic path.

Since the beginning of my academic life as physicist and lecturer in sciences of (social) communication, the classical work of Shannon and Weaver (1949) attracted my attention that was mostly a constant bibliographical reference in the national and international scientific production of my academic area. Often the investigated issue did not justify the inclusion of the reference, which was a mere

entry of a bibliographical digest. Other times, when there were better reasons to do it, the topic was dispatched, yet recognizing the merit of having opened a road, by saying almost a priori: that this road was already exhausted; that the mathematical model of communication was only valid for engineering; that it was unable to realize the *meaning* – which Shannon had explicitly excluded–; and that it led us to a *"transporter" model* of communication (both essential aspects for the study of social communication). No many other arguments were given –speaking in past by politeness–, probably due to ignorance of the mathematical apparatus, of the precedents, of their varying evolution or, merely, because it was not any more the fashion. In Thomas Kuhn's terminology (Kuhn, 1971), the paradigm had already been changed. To be elegant, and especially after my long dedication to university management, I will not give names or works, even exceptions, although there are very distinguished ones. I think this description is enough to portray the interest in formalization in my academic field, when I began to walk through it.

But I had the intuition that it should not be that way. Thus with the baggage of my scientific training and encouraged by one of my first "masters", Santiago Montes (*rip*), I intended through crosscurrents to investigate the relationships between that mathematical information concept and meaning. I tried to demonstrate that behind Shannon and Weaver there was not necessarily a *transporter* model of communication: the model of these authors can also be understood as a *sharing model*. Since then and until the present work I have gone covering different stages of a study that up to now I do not want to consider ended. It is true that I was busy with other applied research works and that with the elapsing of the years my dedication has been less intensive, practically absent in the last eight years, since I was almost exclusively devoted to my executive commitment with the Rey Juan Carlos University, as well as to the creation and consolidation of a Research Group in Communication, Society and Culture (GICOMSOC). Nevertheless, I have always been prepared somehow, though tangentially,

to return to what I considered my principal work line. I believe this work gave some fruits, perhaps modest; the last of them (though some years went by) is the one that I will expose here.

But returning to fashions, the theory of Shannon and Weaver[1] was for a long time more than what is now. The concepts and the model derived from it deluged various fields of knowledge. As I wrote in the introduction to a previous work:

"There was a time in which the Mathematical Theory of Communication was considered as a "fruitful domain" where numerous and various scientific disciplines could meet. The scientific community received Shannon's theory enthusiastically, and used it as model to approach problems in so distant fields as Linguistics, Aesthetics or Theology. After that initial enthusiasm, perhaps exaggerated, the disillusion and the lack of interest for what the Shannon's theory actually provides appeared. The absence or dearth of valid results, the wrong or rushed applications, the lack of rigor and perhaps the impatience, caused the gradual losing of attention to those disciplines.

(…)

A prudential time has already been elapsed since its apparent depletion (the beginning of the interdisciplinary declivity of the mathematical theory of communication can arbitrarily be located in the International Colloquium of Royaumont, 1965), going through a period in which information theory has been developed without the distortion of an excessive attention of foreign disciplines. At the same time, other processes have been explored in the formal study of human communication. Therefore, it is now useful and convenient returning to outline the study of human communication from the

---

[1] To simplify this introduction, I only mention these authors, although before, meanwhile and after Shannon & Weaver, there were actually many scientists who have struggled to develop this trend. Just to mention some of them: Hartley, Nyquist, Küpfmüller, Ashby, Wiener, von Bertalanffy, Bar-Hillel, Carnap, Jakobson, Hintikka, etc.

perspective of the Mathematical Theory of Communication." (Pérez-Amat, 1990, pp. VI-VII)

As I had indicated in the previous reference, the beginning of the academic forgetfulness of the Mathematical Theory of Communication may be located in the International Colloquium of Royaumont in 1965 (Couffignal, 1965). But a certain and brief recovery of the interdisciplinary interest happened in the seminars celebrated in the Faculty of Information Sciences at the University Complutense de Madrid, sponsored by the Spanish General Direction of Scientific and Technical Research (DGICYT), during the academic years of 1992/93 and 1993/94 (Cafaffarel, 1996), in which I had the honour of participating. Meanwhile, my work was developed in a certain –though productive– loneliness, as it was proved by the fact that in these years, beyond the mentioned seminars, I only found two works similar to mine in the academic area of social communication sciences in Spain. I am referring to the Doctoral Thesis defended at the University Complutense by Juan Miguel Aguado (1998), who is now an associate professor at the University of Murcia, as well as to the investigation on "Chaos and Communication" by Ismael Roldán (1999) at the University of Seville.

This framework of scientific and academic "life" places the present investigation, which is related to: 1) my comprehension of *meaning*, or *sense* –if preferred–; 2) the semiotic-cybernetics construction of a model of user communication and of the communication itself; 3) the elaboration of a formal model for the analysis of the *semantic-pragmatic* information.

To conclude this introduction, I will give myself the permission to quote "something" only obliquely related to what I have been arguing, but which underlines the idea that there is a relationship between mathematics and communication, between mathematics and journalism. The author of the quote is the American mathematician John Allen Paulos, who wrote:

"It's time to let the secret out: Mathematics is not primarily a matter of plugging numbers into formulas and performing rote computations. It is a way of thinking and questioning that may be unfamiliar to many of us, but is available to almost all of us.

As we will see, "Number stories" complement, deepen, and regularly undermine "people stories." The notions of probability and randomness can enhance articles on crime, health risks, or other societal obsessions. Logic and self-reference may help to clarify the hazards of celebrity, media spin control, and reportorial involvement in the news. Business finance, the multiplication principle, and simple arithmetic point up consumer fallacies, electoral tricks, and sport myths. Chaos and nonlinear dynamics suggest how difficult and frequently worthless economic and environmental predictions are. And mathematically pertinent notions from philosophy and psychology provide perspective on a variety of public issues. All these ideas give us a revealing, albeit oblique, slant on the traditional Who, What, Where, When, Why and How of the journalist's craft." (Paulos, 2001, p.11).

1. A SEMANTIC - PRAGMATIC MODEL FOR INFORMATION ANALYSIS

"The fact that one cannot interpret a discourse unless one can follow it, suggests that an algorithm which could interpret an arbitrary discourse would have to be "smart" enough to survey all the possible rational, semirational and not-so-far-from-rational-to-be-somehow intelligible discourses that physically possible creatures could physically possible construct." (Putnam, 2000, p. 87)

In previous works (Pérez-Amat, 1996), I outlined the theoretical path –basically *epistemological*– that allowed me: 1) to clarify the relationships between *information* and *meaning*; 2) to establish some solid bases that enable the construction of a (mathematical) theory of information, which would take into account the actual role played in human communication by meaning –a certain meaning notion– or sense –if

preferred–. Part of that effort is already done, and it is only left here the gathering of pieces in the theoretical "puzzle", as well as the formulation of a proposal that could lead towards an analytical *semantic-pragmatic* model of information.

Before undertaking the exposition of this stage in our research, it is convenient to recall some of the conclusions reached so far concerning the basic concepts determining the aims of the work. In its most general sense, *information* can be understood as that reducing uncertainty, whatever its origin might be. In the field of human communication, the *information* only makes sense if it is framed within a *finished* or intentional *action*. Thus, for the study of human communication, *information* has to be defined as that which decreases uncertainty with regard to some kind or model of intentionality. As far as it is concerned, meaning is a problematic notion, whose study should be outlined from the empirical perspective of the use of language. In order to do it, not only a satisfactory model of the user of language is needed, but a satisfactory communication model as well.

1.1. [Belis' Quantitative-Qualitative Information Approach][2]

A good starting point to search for the "satisfaction" of such models is the Mariana Belis' notion of *semantic source* (Belis, 1975). But before and due to the requirements of the ulterior analysis, let us see an extension of Shannon's amount of information, which in a certain epistemological frame may be characterized as *pragmatic information* (Pérez-Amat, 1990, pp. 448-464). However, in order to avoid sterile discussions, I will designate it as their authors, M. Belis and S. Guiaşu, *quantitative-qualitative* measure of the information (Belis and Giaşu, 1968). This deals with the information measure, whose initial cybernetic pose can be summarized in the following terms:

"The cybernetics analogy between man and machine consists (…) in the fact that both are control systems. This means that information is transmitted and processed towards an objective, and control signals should be efficient with respect to it. All the activity of a cybernetic system (biological or technological) is aimed at the achievement of an objective. Thus, the system must have a criterion of qualitative discrimination of signals (….)
The cybernetic criterion for discrimination of signals is represented through relevancy, meaning or utility of information (…) with respect to the objective.
The occurrence of an event reduces a double uncertainty: the quantitative uncertainty concerning its probability occurrence, and the qualitative one concerning its utility for the achievement of the objective." (Ibid, p. 353)

Let there be a finite set of events E_1, E_2,…., E_n, with their corresponding –objective– probabilities of occurrence p_1, p_2, …, p_n, and their corresponding –subjective– utilities with respect to a given objective u_1, u_2,…, u_n.. The occurrence of an individual event provides an information I, whose measure is defined as a function of probability and utility: $I = I(u,p)$. This function must satisfy two obvious properties. *First*: let E and F be two independent events, whose probabilities of occurrence are p and q, respectively. Since the probability of the event set ($E \cap F$) is equal to the product of the probabilities of the independent events ($p \cdot q$), if the events are undistinguished with respect to the utility, that is to say, if all events, including the event set, have the same utility u, then the information of the event set must be equal to the sum of the information of the independent events:

$$I(u, p \cdot q) = I(u, p) + I(u, q)$$

Second: if the utility u of an event is increased in a given value, $\lambda \cdot u$ ($\lambda \geq 0$), then the information provided by that event must be increased in the same value:

$$I(\lambda \cdot u, p) = \lambda \cdot I(u, p)$$

where it can be deduced that,

$\forall\, u \geq 0$ *and* $0 \leq p \leq 1,$ $I\,(u, p) = u \cdot I\,(1, p)$

If in the expression of the first property that must satisfy this quantitative-qualitative measure of information, I make the utility equal to unit, we obtain:

$$I(1, p \cdot q) = I(1, p) + I(1, q)$$

And if I make the function $I(1, p)$ logarithmically depending on probability,

$$I(1, p) = F(\log p)$$

then,

$$F(\log p \cdot q) = F(\log p) + F(\log q)$$

And if this equation takes into account the property of the logarithm of a product, then it will be transformed in:

$$F\,(\log p + \log q) = F\,(\log p) + F\,(\log q)$$

whose unique continuous solution is:

$$F\,(\log p) = a \log p$$

where "a" is an arbitrary constant. Thus, the expression deduced from the second property, which this measure must satisfy, will now take the form:

$$I\,(u, p) = u \cdot I\,(1, p) = u\, F\,(\log p) = a \cdot u \log p$$

And if I now consider that a is an arbitrary constant, it can be made that $a = -K$, through which we can obtain the desired measure of information provided by the occurrence of an individual event:

$$I\,(u, p) = -\,K \cdot u \log p$$

For extension, the mean quantity of information, provided by a set of independent events, can also be measured quantitatively-qualitatively as:

$$I(u_1, u_2, ..., u_n, p_1, p_2, ..., p_n) = \sum_{i=1}^{n} p_i I(u_i, p_i) =$$

$$= -K \sum_{i=1}^{n} u_i p_i \log(u_i, p_i)$$

This quantitative-qualitative measure of information, when it is applied to cases of non existing intentionality, or it is not taken into account (that is to say, when the utility is the same and equal to one for every event), is converted into Shannon's measure of entropy, as it can be easily proven.

Let us consider now the already mentioned notion of *semantic source* from Mariana Belis. A semantic source of information is defined by a finite set of symbols, designated as *dictionary*, on which a finite set of combining rules applies. These rules, designated as *grammar*, generate a finite set of structures (sequences or chains of symbols) that constitutes the *language* of the source. Till now, insofar as grammar is translated by a distribution of conditional probabilities, there is nothing different from Shannon's source of information and from the narrowly related grammar of finite states. According to Belis, a model as the previous one is insufficient for her purposes, because the mere application of the finite grammar can syntactically produce correct sentences –debatable and debated issue, out of the scope of this work–, but meaningless. Therefore, to avoid that the source of information could generate sentences without sense –a debatable issue as well–, Belis introduces the concept of *coherence domain*, which she defines –from the semantic point of view– as "(…) a set of propositions bounded by a *common topic*, for which they are more or less relevant" (Belis, 1975, p. 264). Mathematically, a *coherence domain* is a fuzzy (sub) set, whose characteristic function, designated as *semantic function*, represents the degree of relevancy of each structure (sentence) of the source language with respect to the common topic.

Let there be s_1, s_2, …, s_m symbols –which are also called words– of the dictionary S_m of the information semantic source. Let there be o_1, o_2, …, o_n, the structures or sentences produced by application of the rules of the grammar G on the symbols of S_m; sentences that constitute the language L_n of the source. A coherence domain d_j in L_n is a fuzzy subset of L_n in which the semantic function $f(o_i)$ takes a value between 0 and 1 for each sentence o_i. The value $f(o_i) = 0$ represents the absence of meaning of o_i in the domain d_j, while the value $f(o_i) = 1$ indicates that the sentence o_i has a "great meaning" in d_j.

What does it mean that a sentence holds a "great meaning" in a given coherence domain? According to the aforementioned definitions, it means that the sentence belongs clearly to the set of propositions that

constitute such coherence domain; that it is "absolutely" pertinent or totally relevant for the *common topic* that characterizes the domain. But what does it mean *common topic*? The expression is too ambiguous to leave it without further explanation. Belis did not clarify it more than what has been mentioned before, thus tingeing with ambiguity not only the definition of coherence domain, but also the set of her theoretical position as a whole. I will try to give that explanation, or –better said– a possible precise interpretation of the concept of coherence domain, when I finish exploring Belis' theory.

Thus, the model of semantic source of information also includes a finite set *D* of coherence domains, d_1, d_2, …, d_p,, so that a same structure or sentence of the source language has associated a semantic function that takes different values, though not necessarily different, in the different domains of *D*. That is to say, the meaning of a sentence depends on the coherence domain in which it is produced, which causes that the concept of coherence domain will be somehow equivalent to the *context* notion. Each semantic source holds a specific preference order of the coherence domains, which is represented by a finite set of real numbers K_j ($0 \le K_j \le 1$) that are assigned to each domain d_j of the source. These numbers are designated by Belis as *preference coefficients*.

This formal model of semantic source of information is adapted to describe any semantic system whose operating mode is either deterministic or not, as producer of symbols or transmitter of messages. Within the second case, Mariana Belis is especially interested by semantic systems that operate in a mode that she designates as "stochastic-controlled" ones, namely, if the election of symbols is randomly done –Markov process–, although controlled by coherence domains, so that the successive sentences, produced by the system, have high values of the semantic function in the initial coherence domain. All this occurs in case the semantic system is a source of properly called information, that is to say, when the system acts as emitter. When the semantic system acts as receiver, the model of semantic source continues being valid for the description of its operative

process: the system recognizes and decodes signals received by comparing with the symbols and structures that, in relation to the set of coherence domains, has been stored in an operative memory.

In order to produce communication between two semantic systems as those conceived by Belis, the situation must satisfy two initial conditions, although partially. In such case the quantity of transmitted information reduces in proportion to the degree of incompleteness. These conditions are:

1.   Both systems should hold the same *dictionary* and the same *grammar*, which implies that they should produce the same *sentences*, that is to say, they should have the same *language*.

2.   There must be a common *coherence domain* to which both systems assign approximately the same *preference coefficient*.

Let there be two semantic systems A^1 and A^2:

$$A^1 = \{S^1,\ G^1,\ L^1,\ D^1,\ K^1\}$$
$$A^2 = \{S^2,\ G^2,\ L^2,\ D^2,\ K^2\}$$

If the aforementioned initial communicative conditions are satisfied, then:

$$S^1 = S^2,\ \ G^1 = G^2,\ \ L^1 = L^2$$
$$K_0^{\ 1} \approx K_0^{\ 2}$$

I will also assume that the successive transmission of structures can be treated as a Markov process[3] of order *h*, and that the communication process is stationary, although in finite temporary intervals. Normally, communication between systems with learning ability is not stationary, and we are especially interested in these systems. Therefore, the stationary of the communication process should be assumed only for finite temporary intervals. If we consider each source in absence of the other, their *characteristic entropies*, according to

---

[3] Here, the dependence of the Markov's process, unlike in Shannon's theory, is not established between signals or symbols, but between structures or sentences, which implies that the determination of conditional probabilities is much more difficult.

Belis' *quantitative-qualitative* measure of information, will be expressed as:

$$A^1 : H^1 = -K_0^1 \sum_{i=1}^{n} f_0^1(o_i)p^1(o_i \mid o_h)\log p^1(o_i \mid o_h)$$

$$A^2 : H^2 = -K_0^2 \sum_{i=1}^{n} f_0^2(o_i)p^2(o_i \mid o_h)\log p^2(o_i \mid o_h)$$

where $p^1(o_i|o_h)$ and $p^2(o_i|o_h)$ are the conditional probabilities of the structure or sentence o_i when h structures have already been produced by the source A^1 and A^2 respectively.

At the beginning of the communicative interaction between the two semantic systems, each one of them "ignores" what the characteristic entropy of the other is. But they do not "ignore" each other. The (subjective) knowledge that a semantic system has about the other system is represented by the *reflected entropy*, namely, the entropy of a source seen by the other. Thus, the reflected entropy of A^1 (in A^2) is:

$$H^{1,2} = -K_0^{1,2} \sum_{i=1}^{n} f_0^{1,2}(o_i)p^{1,2}(o_i \mid o_h)\log p^{1,2}(o_i \mid o_h)$$

and A^2 (in A^1):

$$H^{2,1} = -K_0^{2,1} \sum_{i=1}^{n} f_0^{2,1}(o_i)p^{2,1}(o_i \mid o_h)\log p^{2,1}(o_i \mid o_h)$$

where the second super index indicates that the values of the preference coefficient, the semantic function and the conditional probability are "subjectively" assigned by the system, labelled with such super-index, to the system, labelled with the first super-index. The difference between those "subjective" values and the objective values defining the initial state of the system, that is, between the reflected entropy and the characteristic entropy, depends on the initial knowledge of a system with respect to the other. This knowledge normally increases during the communication, since a process of *pattern recognition* is produced in it. Through these processes, one of the systems "(…) discovers the real probability distribution of the other one (…) as well as the values of its preference coefficient and of their semantic functions" (Belis, 1975, p. 268).

Simultaneously with that pattern recognition, other process is produced in communication. Through it, the characteristic entropy of a source is modified by the presence of the other one. It is the *mutual influence* process, represented by the *conditional entropy*.

Let $p^{1|2}(o_i|o_h^1, o_h^2)$ be the conditional probability that source A^1 transmits the sentence o_i when previously h sentences have been transmitted by A^1 and A^2; let $f_0^{1|2}(o_i)$ be a semantic function corresponding to the sentence o_i in the coherence domain d_0 of system A^1, modified by the presence of the system A^2; and let $K_0^{1|2}$ be a preference coefficient assigned to the domain d_0 for the system A^1 in presence of the system A^2. Then, when each system has previously transmitted h structures or successive sentences, the conditional entropy of A^1 in presence of A^2 is given by:

$$H_h^{1|2} = -K_0^{1|2} \sum_{i=1}^{n} f_0^{1|2}(o_i)p^{1|2}(o_i \mid o_h^1, o_h^2)\log p^{1|2}(o_i \mid o_h^1, o_h^2)$$

and the conditional entropy of A^2 in presence of A^1:

$$H_h^{2|1} = -K_0^{2|1} \sum_{i=1}^{n} f_0^{2|1}(o_i)p^{2|1}(o_i \mid o_h^1, o_h^2)\log p^{2|1}(o_i \mid o_h^1, o_h^2)$$

Thus, as the conditional entropy has been defined, it is applicable to processes of dialogic communication, in which each system acts alternatively as transmitter and as receiver. Belis, although not explicitly, only refers to this kind of communication, what does not mean the unsuitability of the model for unidirectional communication processes, in which a system acts always as transmitter and the other as receiver, neither that in such processes the influence disappears, although it is no longer mutual. The influence of the transmitter A^1 over the receiver A^2 disappears –when a semantic system only acts as receiver, it makes no sense to speak about any of the considered entropy types–, but the possibility of influence of the receiver over the transmitter is maintained, though certainly only in its qualitative aspect (i.e., semantic functions and preference coefficients), not in its quantitative aspect (i.e., conditional probabilities distribution).

Since the mutual influence is a process that elapses simultaneously within transmission, the value of the conditional entropy of A^1

and/or A^2 varies continually during communication. Consequently, the process of pattern recognition, which is simultaneous and is interrelated with the mutual influence, causes that the reflected entropy evolves coupled with the conditional entropy, i.e., it continually reflects –starting from the initial characteristic entropy– the successive conditional entropies of each system. On the other hand, pattern recognition produces the successive approximation of the value of reflected entropy to that of conditional entropy. The maximum value of this approximation depends on the situation and degree of cooperation or antagonism in which the communication is carried out.

The amount of transmitted information from a semantic system to other is measured by the difference between the conditional entropy and the reflected entropy. This quantity will be labelled here as *semantic transinformation*[4], which –according to Mariana Belis– "(…) represents the difference between the authentic *informational content* of a source and the knowledge that the other source has about it" (Belis, 1975, p. 269). Thus, the semantic transinformation from A^1 to A^2 is:

$$T_{1,2} = H^{1|2} - H^{1,2}$$

and the one from A^2 to A^1:

$$T_{2,1} = H^{2|1} - H^{2,1}$$

Now, if the conditional entropy and the reflected entropy vary continuously during communication, the semantic transinformation will do as well, reducing its value as the process advances[5]. Thus: when can the semantic transinformation be computed? What is easy, and the only possible way in cases of unidirectional communication, is to make it before the beginning of transmission (here it cannot be properly spoken about semantic transinformation) and after it

finishes. But this is not useful to observe the double evolution of the semantic transinformation in cases of dialogic communication, even to obtain the total amount of transmitted information in both directions for such cases. However as we defined conditional and reflected entropy, there is actually no problem, because communication proceeds operatively in the model, in *h* successive stages, where it is possible to compute first the reflected entropy value and afterwards the semantic transinformation one. Those stages correspond to the successive cycles in which the same system acts as a transmitter and as a receiver. The double distribution of the values of semantic transinformation allows us to observe the evolution of the double process of pattern recognition. The total amount of (semantic) information being transmitted by each system is obtained by adding the corresponding values to the semantic transinformation in each stage of the communication process.

The parallel evolution of the mutual influence process can be observed through the values taken by the *malleability index* in the *h* stages of the communication process. This index is obtained subtracting the successive values of the conditional entropy from the value of the initial characteristic entropy of each system. The final value of this index, after the last stage of the process, is properly known as *malleability* of the semantic system, and permits to distinguish two classes of sources or systems: *weak systems*, with a large malleability, i.e., they are easily influenced; and *strong systems*, with a small malleability, i.e., they are hardly influenced.

1.2. [Limits of Belis' approach]

This theory of "semantic communication" from Mariana Belis, clearly "informational" and cybernetically oriented, fits with the cybernetic description of semiosis in symbol processing systems that I examined in my work of 1990 (Pérez-Amat, 1990, pp. 320ff), insofar as the sets and functions {*S, G, L, D, f(o), K(d)*} defining the semantic source of information can be understood as a formal equivalent – "informationally" operative– of the cognitive adaptive map and of the variable intentional

---

[4] The notion of semantic transinformation agrees –with particular variations of the considered semantic aspects– with the classical definition of mutual information or transinformation in the mathematical theory of communication and/or information (See Abramson, 1963, pp. 124-; Nauta, 1972, p. 194 ; Yaglom, 1969, p.82-).

[5] In case of dialogic communication, usually, each transinformation (from A^1 to A^2 and from A^2 to A^1) gradually decreases its values in a different measure.

state of the interpreter-systems with symbol processing ability. Thus, I believe that the cybernetic semiotics can provide the general theoretical basis, needed by a model so concisely mathematical, as that of Mariana Belis, in order to transcend social-anthropologically the purely formal area of artificial languages. We believe it, even though, if Doede Nauta (main co-author of such cybernetic semiotics) had been asked, he would not have admitted the semantic nature of such communication model. For Nauta, the information at stake in this model would not be (meta)-semantic, but semiotic, and more specifically, discursive information (Nauta, 1972, pp. 63ff)[6]. However, since this discrepancy on information nature is based on the weakest component of the theoretical construction of Nauta (Pérez-Amat, 1990, pp. 387-392; Llorens, 1974), it has little effect to the support offered to Belis' model by such theory –as a general theoretical framework– or by any other able to describe cohesively the semiosis process in cybernetic and systemic terms.

Now it is time to ask two successive questions: is the Belis' theory sufficiently consistent? And, if this is the case, does it totally satisfy our expectations? The ambiguity of the coherence domain concept that I pointed out before, seems to indicate that both answers tend to be negative. Let us analyze this issue more in detail.

According to the formal definition of the coherence domain concept, the semantic function measures the greater or smaller relevancy of the different sentences with respect to a "common topic". That is, the relevancy is related to the specific core topic characterizing each coherence domains of a semantic information source that "controls" the stochastic production of sentences. Consequently, the "common topic" constitutes the central axis, assumed to articulate a whole communication process, so that for every communication –insofar as closed process–, a constraint of containing only thematic and textual unit has to be presupposed. Even though, it is a dialogue or a nonsensical discussion. In other words, each thematic unit

of a text determines a complete communication cycle.

Furthermore, the formal definition of coherence domain demands other limitation, since, according to the model of semantic source, each of them holds a finite number of coherence domains. Therefore, a semantic source will only be able to communicate intentionally about a finite number of "topics". On the other hand, Belis explicitly admits that his concept of coherence domain is in a certain way equivalent to the notion of context. She does not tell us which class of context, but, as seen before, it can only be a thematically limited context, of semantic-linguistic nature.

Accepting a certain dose of reductionism –a drawback we have to live with, though only in a certain degree–, if the operative definition of the concept of coherence domain and the stochastic-controlled proceeding mode of semantic systems seemed to assure the sense of all possible sentences we might not have much to object. But that is not the case, since they can only assure the significance, i.e., the relevancy of each sentence regarding a central common topic. However, as said before, the need of that double warranty is just what justifies the incorporation of the coherence domain concept.

It might be the case that Belis had assumed the sense of all sentences which could be produced in his model of "semantic communication". That seems to be pointed out when, in the definition of semantic source, he includes a finite set L_n constituted by the unique n structures or sentences that can been generated by the source in any circumstance. But in such a case, since the n sentences are composed from the m symbols or words of the dictionary through application of a finite set of grammatical rules, Belis should assume a grammar (maybe transformational or of other type), whose recursive rules only produce meaningful sentences. However, it does not happen in the finite states grammar, based on the same Markov process he uses to elaborate his theory. Thus, the theory of the "semantic communication" would have an epistemological hybrid nature, since on the one hand, it makes effective use of a

---

Markovian "grammar" in the textual level –if I am allowed to use the term–, and on the other hand, it would need to presuppose a grammar of other type in the sentence level, which –I think– might be qualified at least as eccentric.

However, we do not believe that Belis has so heterogeneous conception –in spite of the fact that her text offers signs pointing to this direction–, because it also gives samples of the contrary with affirmations as the following:

> *"The mere application of the grammar may generate sentences, which are right from the syntactic point of view but meaningless. We introduced the concept of coherence domain to include the (…) meaning".*[7] *(Belis, 1975, p. 264)*

It seems that the stochastic-controlled proceeding mode not only relates to the successive production of the text sentences, but also to the production of the sequences of symbols or words, i.e., sentences, which send us back to the previous issue about whether the model can guarantee or not the sense of all its possible sentences. In principle, my posture concerning this issue is something pessimistic, because, as I previously indicated, the values of the semantic function only represent the greater or smaller relevancy of the successive sentences with respect to a central topic. Though certainly, if the relevancy is guaranteed –in whatever degree, normally high– with respect to a "common topic", it is also assured that the sentences have some sense in the coherence domain of that "common topic". Indeed, any meaningless sentence cannot be relevant with regard to that "common topic" or to any other one. Consequently, the meaningless sentences remain implicitly excluded from the model by the assignment of a semantic function –semantic value– equal to zero in all the coherence domains of the set *D*.

Nevertheless, this reasoning, perhaps valid for a theoretical strategy of "semantic communication" –as that of Belis' model–, more devoted to the "informational" performance than to the meanings themselves of successive sentences, it is absolutely not suited from a semantic-

linguistic perspective, because it makes no sense that the meaning of a sentence or its lack of meaning be determined by its relevancy or irrelevancy regarding all possible topics to communicate. From the linguistic point of view, Belis' procedure is improper, something like "to build a house starting from the roof", because, in order to decide the degree of relevancy or irrelevancy of a sentence in a domain or given context, it has to be previously known that the sentence actually holds some meaning, and if that is the case, what is(are) that(those) meaning(s) in all the possible contexts wherein it could be given.

In any case, when the merely "informational" objectives of Belis' theory –not so far from mines– has been observed, the inappropriateness that can be linguistically attributed to the model of "semantic communication" is not so fundamental as to reject it totally: as (mathematical) theory of information fulfils its objective of including the semantic aspects satisfactorily.

1.3.   [Fuzzy nature of semantic information]

Now, is that all what can be said about meaning –"informationally" speaking– or is there anything else to analyze in this regard? Let us imagine a very possible case in Belis' model: two sentences whose values of the semantic function will be identical in a same coherence domain, but one of them in its literal denotative sense and the other in a figurative sense. Is the same amount of information transmitted with both sentences? In terms of the textual coherence, evidently yes. But perhaps in terms of the sentences themselves, it is not. Is there not any extra information that one of them transmits with respect to the other; a difference in the uncertainty that they reduce? Yes, I think there is an added uncertainty reduction, some extra information of semantic foundation, which –in opposition to Belis' information– does not have a probabilistic character under the intentional control of a fuzzy variable, rather its nature is directly "fuzzy". It is the information that I designated as *structural* in a previous work (Pérez-Amat, 1990, pp. 30-38). Thus, when a semantic system produces a

---

[7] NT: Translated from the Spanish version of Pérez-Amat.

sentence, the qualitative (semantic) information being transmitted has two components: first, the one studied by Belis, which is related to the degree of textual coherence, and another related to the fuzzy or diffuse nature of the meaning itself. The measurement of this last component and its integration in the model of "semantic communication" are tasks that exhibit many difficulties, but there are fortunately some "tools" that might facilitate our attempt to do it, though our essay will be no more than incomplete.

Some lines above, I referred –without going into details– to the "fuzzy or diffuse nature of meaning", but this expression requires some clarifying explanations. First, we should remark that asserting this expression makes no sense if the language of the considered semantic source is artificial, and therefore holds univocal semantic rules. In such case, asserting the existence of a structural component of information is also meaningless. The diffuse nature of meaning can only be asserted for diffuse artificial languages[8] or, in our concern, for natural languages. According to Goguen, natural language has a property known as *robustness* (Gouguen, 1975) through which their users have the "ability of answering without program modification to slightly disturbed or incomplete defined situations". In other words, correct understanding is not hindered by diffuse situations as those of syntactic imperfection of sentences: for the comprehension of language, it is not necessary the good syntactic formation of sentences.

Now then, the diffuse character does not only relates to the language with regard to the exhibited robustness against situations of syntactic indefiniteness, but also to its semantic *vagueness*, namely the "meaning" of words is diffuse, or in other words, accurate usage limits can not be traced in most cases. Furthermore, this semantic vagueness, this "flexibility" of the "meaning" of words, or of its sense –if preferred–, mediates decisively in the linguistic comprehension, making it easier by means of "semantic short-cuts", which requires smaller effort. But, what is even more important, it makes possible that the

comprehension is produced according to the principle of maximum meaningfulness of sentences in a given context. Concerning this issue, it should be remarked that the mathematical model of Belis attains –not explicitly- this semantic optimization principle of Goguen, since the control, executed by the coherence domains on the stochastic emission and reception of sentences, is done by means of selecting those whose semantic function has the largest value in the considered domain. That is to say, those sentences having an optimal significance for the context, configured by a specific coherence domain, are selected.

Regarding the *fuzzy* or diffuse conception of lexical meaning, I have already indicated (Pérez-Amat, 1990, pp. 205ff) some partial works aiming at that direction from various theoretical perspectives: the "prototypical categories" from Rosch (1978), the Labov's experiments on "cups", "mugs", "bowls" and "vases" (Labov, 1973), the "semantic hedges" from Lakoff (1972), the "linguistic variables" of Zadeh's Quantitative Semantics[9], etc. I will now tackle a formalization of the "fuzzy meaning" that by its mathematical character will have a more general application and we might integrate within the model of "semantic communication" from Mariana Belis.

1.4. [A quantitative approach to fuzzy semantic information]

In previous works I agreed with Hilary Putnam (1983) in his appraisal of the fact that, yet being accord with W. V. O. Quine (1953, 1969, 1973, 1964) concerning the non-existence of a thing or defined object called meaning, the meaning processes can be studied through the factors that get involved in the transmission of the normal uses of language. This transmission is produced by communicating a set of prototypical categories, namely, the central or core facts that, as empirical hypothesis, define the "meaning". In such works, I asserted that if

---

[8] See, for example, (Zadeh, 1971)

[9] A "linguistic variable" is a variable whose values are words or sentences of natural or artificial language. For example, "age" is a linguistic variable if its values are: "young", "very young", "more or less young", etc. Every term of a "linguistic variable" is a label of a fuzzy subset in a discourse universe (Zadeh 1971, 1975).

there are some central facts to teach the normal use of words, there must also be some other more or less peripheral facts, whose knowledge is necessary to communicate in less normal circumstances, that is, in contexts that are far away from the "denotative conceptual norm".

Now, if –in a much broader context– the "meaning" of a term is considered as formed by a set of prototypical or core categories, to which a value 1 is assigned, and a (finite) series of sets of peripheral categories –some characteristics can belong to several sets of a same meaning-, to which values between 0 and 1 are assigned in function of the greater or smaller distance to the core or centre, then we have a definition of "fuzzy meaning". Thus "meaning" is represented through a fuzzy subset of the partition set of the speech universe, whose characteristic function, µ, measures the weight representing the contribution of each set of facts to the global "meaning" of the term.

I will return now to Belis' model of semantic source. If I want to integrate it with the notion of "fuzzy meaning", I should have a *semantic function*[10], defined on the symbols or words of the source dictionary and representing the degree of "indefiniteness" or ambiguity of its "meaning". The first way for obtaining such function, which I will refer by $g(s)$, is by means of calculating the indefiniteness or mean ambiguity of the diffuse subset representing the "meaning" of each symbol or word s:

$$g(s) = \frac{1}{q}\sum_{t=1}^{q}\mu_t$$

The previous function, whose values fluctuate between 0 and 1, can be interpreted as the characteristic function of the symbols subsets making up the structures or sentences of the source language. Each sentence of the source language L_n is formed by a sequence of symbols of the dictionary S_m; the set of symbols taking part in each sequence can be considered as a diffuse subset of S_m, in which the characteristic function, $g(s)$, is a semantic function representing the indefiniteness or previous

ambiguity provided to the sequence by the "meaning" of each symbol. When the source transmits or produces a sentence in a given coherence domain, each symbol or word attains a sense (supposedly accurate), so that an uncertainty reduction is achieved, i.e., a structural information is transmitted, which can be measured by the non probabilistic entropy defined by Luca and Termini (Luca, 1972). Thus, the amount of transmitted (semantic) structural information for each sentence o_i, generated by a source in a communication process, will be given by the following equation:

$$\psi_i = B\sum_{t=1}^{r} S[g(s_t)]$$

where $S(g)$ is the function of Shannon:

$$S(g) = -g \cdot \ln g - (1-g) \cdot \ln(1-g)$$

and B is a constant which, in case of being a normalized entropy and substituting the base e of Napierian logarithms in Shannon's function by 2, has the value:

$$B = \frac{1}{r\ln 2}$$

Since h sentences will be transmitted in each direction in the h steps of the complete communication process, then the final semantic transinformation will be given by:

$$T_{1,2} = \sum_{i=1}^{h}\psi_i^1 + H^{1|2} - H^{1,2}$$

$$T_{2,1} = \sum_{i=1}^{h}\psi_i^2 + H^{2|1} - H^{2,1}$$

Now then, this mode of facing the problem is not completely satisfactory since the structural information transmitted by each sentence is always the same, independently of the coherence domain wherein it is produced. To solve this insufficiency I have to proceed avoiding one of the given steps, namely that of working with the average indefiniteness of symbols. If the coherence domain, preferred by the source, controls textually the stochastic transmission of sentences, it must also control its "meaning", that is, it must determine the set of more or less central characteristics through which

---

[10] For Belis' semantic function, I propose the more appropriate name of *textual function*.

each symbol contributes in the "meaning" of that sentence. Thus, it is no longer the average function *g*, but *µ* instead, which plays the semantic function roll, so that the quantity of structural information transmitted by each sentence o_i in a given domain d_j is given – according to Luca and Termini's normalized computation– by the following equation:

$$\psi_{ji} = B\sum_{t=1}^{r} S(\mu_t)$$

This type of measuring structural information is not context independent, since each sentence provides a quantity of semantic transinformation that –in principle– does not have to be the same in every coherence domain where it could be produced. Furthermore, counting with function µ –insofar as determined by empirical hypothesis–, we are now in conditions of assuming that the processes of pattern recognition and of mutual influence also relate to semantic definition or indefiniteness. So that the structural component of information, the entropy *ψ*, is now added, with the variations imposed by such processes, to the conditional entropy and to the reflected entropy of Belis in each stage of the communication process, constituting thus the new conditional and reflected entropies, whose difference(s) provide(s) the quantity(-ies) of semantic transinformation.

References

Abramson, N. (1963). *Information Theory and Coding*, New York: McGraw-Hill.

Aguado Terrón, J. M. (1998). *Fundamentos epistemológicos del paradigma de la complejidad: Información, comunicación y autoorganización*. Doctoral Thesis. Madrid: Departamento de Periodismo III, UCM. This work was reviewed and published as: (2003). *Las bases de la complejidad*. Sevilla: Comunicación Social Ediciones.

Belis, M. (1975). "A theory of semantic communication". In J. Rose y c. Bilciu (eds.), *Modern trends in Cybernetics and Systems*. Vol. 3, Berlin: Springer, pp. 263-271.

Belis, M. and S. Guiaşu. (1968). "A Quantitative-Qualitative Measure of Information in Cybernetics Systems". *IEEE Trans. on Information Theory*, *14(4)*, 593-594.

Caffarel, C. (Ed.). (1996). *El concepto de información en las Ciencias Naturales y en las Ciencias Sociales*. Madrid: UCM.

Couffignal, M. L. (dir.). (1965). *Le Concept de l'information dans la science contemporaine, Les Cahiers do Royaumont*, Paris : Gautier-Villars.

Feyerabend, P. K. (1975). *Against method: outline of an anarchistic theory of knowledge*. London: NLB.

Goguen, J. A. (1975). "On fuzzy robots planning". In Zadeh, L. A., Fu, K. S., Tanaka, K. & M. Shimura (Eds.), *Fuzzy sets and their applications to cognitive and decision processes*. New York: Academic Press, pp. 429-447.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Labov, W. (1973). "The boundaries of words and their meanings". In Bailey, C.J. y R. Shuy, *New ways of analysing variation in English I*, Washington: Georgetown University Press, pp. 340-373.

Lakoff, G. (1972). "Hedges: a study in meaning criteria and the logic of fuzzy concepts". In Hockney, D. Harper, W. y B. Freed (Eds.), *Contemporary research in philosophical logic and linguistic semantics*, Dordrecht: Reidel, pp. 221-271.

Llorens, T. (1974-1975). "Información y semiosis" (I & II). *Teorema, IV(1)*, 55-89; *V(2)*, 213-231.

Luca, A. de & S. Termini. (1972), "A definition of a non probabilistic entropy in the setting of fuzzy sets theory", *Information and Control*, *20(1972)*, 301-312.

Nauta, D. (1972). *The meaning of information*, The Hague: Mouton,

Paulos, J. A. (2001). *Un matemático lee el periódico* (Spanish translation). Barcelona, Círculo de Lectores. [Source: (1997) *A Mathematician Reads the Newspaper*, New York: Random House (Anchor Books)]

Pérez-Amat, R. (1990). *Información y Significado*, Madrid: UCM, 1990.

Pérez-Amat, R. (1996). "Información y entropía". In Caffarel, C. (Ed.) *op. cit.* pp. 31-40.

Putnam, H. (1991), *Representation and reality*. Cambridge, Mass.: MIT Press [*Representación y Realidad. Un balance crítico del funcionalismo*, Barcelona, Gedisa, 2000]

Putnam ,H. (1977). "Is semantic possible?" In S.P.Schwarts (Ed.), *Naming, necessity, and natural kinds*. Ithaca, NY: Cornell Univ.Press ["¿Es posible la semántica?", Cuadernos de Crítica, Nº 21, México, UNAM, 1983]

Quine, W. V. O. (1953). *From a Logical Point of View*. Cambridge, MA: Harvard University Press.

— (1964). *Word and Object*. Cambridge, MA: The MIT Press.

— (1969). *Ontological Relativity and Other Essays*. New York: Columbia Univ. Press

— (1973). *The Roots of Reference*, LaSalle, Illinois: Open Court Publishing.

Roldán Serrano, I. (1999). *Caos y Comunicación. La teoría del caos y la comunicación humana*, Sevilla, Mergablum,

Rosch, E. (1979). "Principles of categorization". In Rosch, E. y B. Lloyd (Eds.). *Cognition and Categorization*. Hillsdale (N.J.): Lawrence Erlbaum, pp. 27-48.

Shannon, C. E. & W. Weaver. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press. [*Teoría matemática de la comunicación*. Madrid: Forja, 1981]

Thom, R. (1982). "Mathémalique et theorisalion scienlifique". In APÉRY, et al. *Penser Les Mathematiques: Seminaire De Philosophie Et Mathematiques*. Paris: Seuil. [APÉRY, et al. (1984). *Pensar la Matemática*. Barcelona: Tusquets.]

Yaglom, A.M. & I. M. Yaglom. (1969). *Probabilité et information*, Paris : Dunod.

Zadeh, L. A. (1971). "Quantitative Fuzzy Semantics", *Information Sciences*, *3*, 159-176.

Zadeh, L. A. (1975). "Calculus of fuzzy restriction". In Zadeh, L. A., Fu, K. S., Tanaka, K. & M. Shimura (Eds.), *Fuzzy sets and their applications to cognitive and decision processes*. New York: Academic Press, Appendix.

About the Author

Ricardo Pérez-Amat

(*b.* Almería, Spain, 1952; *d.* Madrid, 2008) studied in the Universities of Granada, Valencia and Complutense de Madrid, where he was awarded first with a B.S. in Physics (astrophysics branch) afterwards in Journalism with special award. Ph.D. in Journalism at the Universidad Complutense, with a Thesis "Information and meaning: towards a semantic theory of information". He became associate professor in 1991 at the Universidad Complutense, where he held several administrative responsibilities. He joined the Universidad Rey Juan Carlos de Madrid in 2000, where he launched three academic degrees in Communications Sciences and was main responsible of the Communication Science Department and Faculty. In 2002, he became full Professor in Journalism at the same University, where he was also responsible for the Master of Agency Journalism in cooperation with EFE news Agency. He also held the vice-rector chair in European convergence and was invited lecturer within doctoral or open programmes at the University of New Mexico, Cervantes Institute of New York (2003), Austral University of Valdivia (Chile) and General Cultural Direction of the Dominican Republic (2007). Pérez-Amat was collaborator of the Ser Radio Channel and the Spanish National Radio as party in radio-gathering programmes.

He authored many research works and an important number of publications in books, scientific journals and specialized press in audiovisual culture. Among them, it may be mentioned: "Información y significado" (Information and Meaning) appendix to the Spanish edition of Shannon, C. and Weaver, W. (1981), *Teoría Matemática de la Comunicación*; *Información y significado* (1990), "Información y entropía" (1996), *Sociedad, integración y televisión en España* (2006), "Telerrealidad o Telemetarrealidad" (2006), *Comunicación, identidad y género* (2008), "Audiences, citizens and politics as a religion: the spanish televisión" (in press). His main research areas were information theory, research methodology and applied communications (mainly in audiovisual contexts). He directed the research group in Communication, Society and Culture at the Communication Science Department, Universidad Rey Juan Carlos.