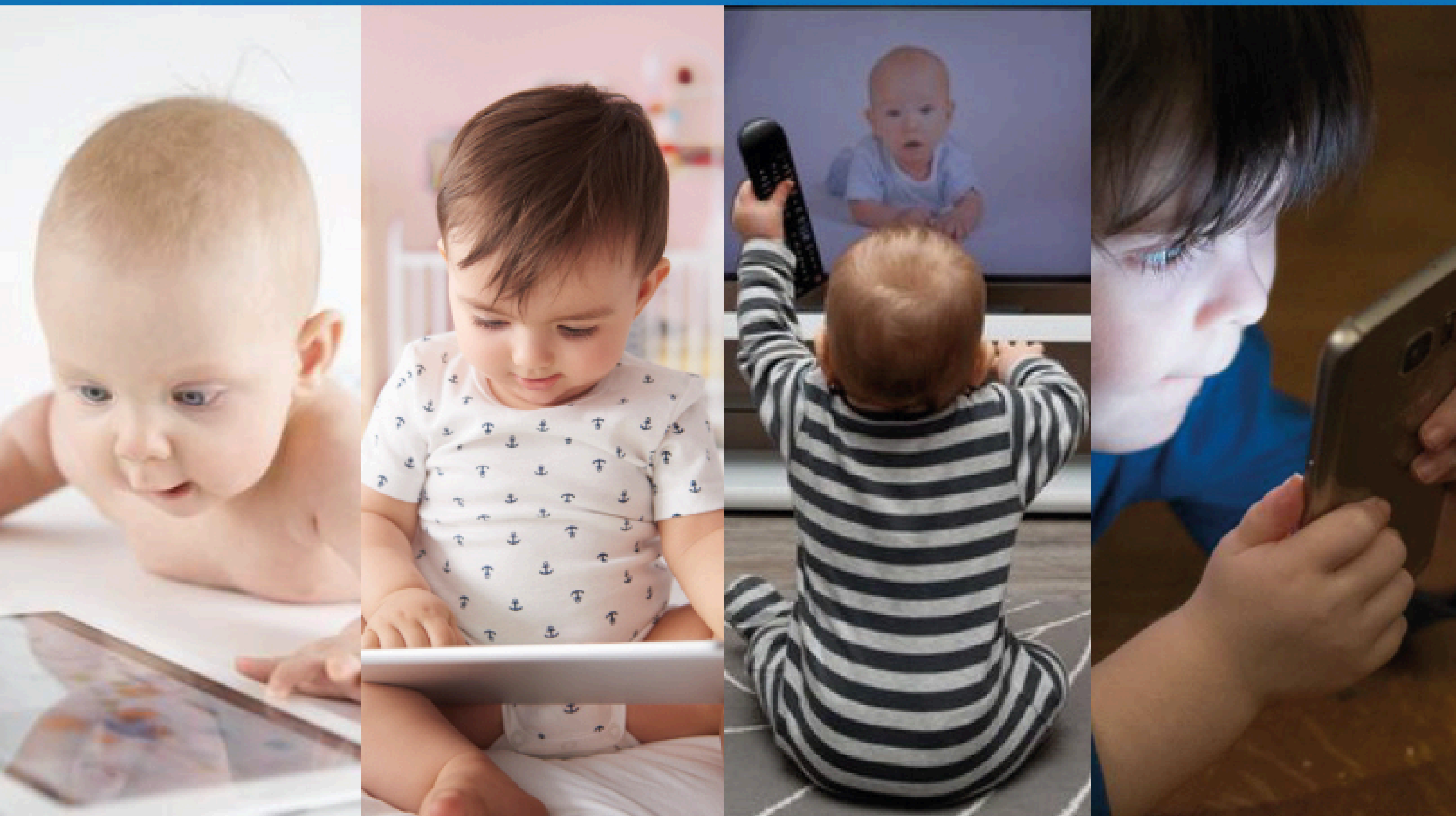
# 5.1 MATHEMATICAL THEORY OF SIGNIFIED



**Benigno Moreno Vidales**



INGENIERÍA NEUROCIENCIA GENÓMICA & INNOVACIÓN TECNOLÓGICA

www.ingit.es

# 5.1 MATHEMATICAL THEORY OF SIGNIFIED

Benigno Moreno Vidales[1]

Barcelona, 15/06/2021

**Index**



**Introduction**

The cornerstone of "*The Mathematical Theory of Communication*", MTC (C. Shannon, 1948), allowed a very important advance in the field of information transmission.

It was approached as a subject of telecommunications engineering, and although N. Wiener also worked on it from the discipline he founded of cybernetics, the fundamental difference was the character of the signal to be transmitted, which in Wiener's case was " analog" or continuous.

In the case of Shannon, the signal is treated in the terms that we now call "digital".
But this is not the only difference, since the term <information> only includes the <physical signal> that is transported, that is transferred to be communicated, and not the semantic content.

C. Shannon intentionally does not refer in his proposal to what we understand in colloquial language by "information"[2], which obviously involves a physical signal (auditory, written, etc.), as well as its <semantic meaning>, and he decides not to contemplate this in your proposal, as we said.

Therefore C. Shannon, at a linguistic level, deals with the signal that the lexicon carries, only in its nature <lexical signifier>, leaving aside the aspect of content that must be associated with the <semantic signified>.
C. Shannon, working at Bell Telephone, and a professor at MIT, adopts that his theory is accompanied by a text by veteran Sloan Foundation adviser Warren Weaver.
In this text, which appears as an introduction to the MTC publication in 1949, extensive reference is made to <semantic meaning>. Its title is: <<*Recent contributions to the mathematical theory of communication*>>.
We will deal with it in Section 1, of this analytical study, with a proposal for a model on the quantification of meaning that Weaber points out throughout his Contributions.

The aspect of the theory of C. Shannon differs from the perspective of N. Wiener, we would say in today's terms, as a model in which one works respectively with a discrete "digital" signal, and another, with an "analog" signal. " continues, as we pointed out above.

---

[1] References at www.ingit.es. Intellectual Property Registry <02 / 2021 / 3606>
[2] "information": in the normal or vulgar sense, which encompasses the "significant signal" and its "semantic meaning" jointly

The crucial difference between the two is that the first works in its fractional or discrete conception, with numerical binary code for very precise signal transmission, in the communication process. The method of digital "atomization" allows to control the precision, accuracy and reliability of the signal, and also being able to pass to an evaluation of the <information>[3] signal, a generalization that the act of communication entails, with the attribute that supposes a freedom of choice of consecutive symbols.
A message within the framework of such evaluation is that of probability theory, in which the symbol to be transmitted must be chosen, which is configured with a zero (0), or a one (1), their set (0-1) is an <information> **bit**.

Beyond Shannon and Wiener, P. Rocchi (2010) recapitulates some 25 information models to date (until 2010), with the most extensive summary and analysis known to date.
And it ends up adopting a semiotic model from linguistics, which "charges" its own toll by ascribing itself solely in terms of that digital model, and that although it makes a generalization towards symbolic systems, it is not complete, although language is for us (and therefore its analytical discipline, semiotics), the symbolic instrument par excellence.

Despite the very important synthesis and choice adopted by Rocchi (Demming & Bell, 2012), it also allows dealing with both the signifier-lexical and the signified-semantic, and on the other hand, it is generalized to the field of telecommunications, using the terminologies of "analog" information, as well as the great leap to information with "digital" technology (from which conversion is made to "analog" when needed; the "bypass" between the two, and in systematic mathematical methodology, is established by methods such as that of Runge-Kutta, 1900), being informational clarifier, and a vital advance in the specialty of telecommunications and computing.

The vision of "meaning" is taken to what has later been designated as "narrow meaning" (the one provided by a word or term, without disambiguation and "purely", an isolated word).
Hinzen and Poeppel[4] (2011), point to the "broad meaning", that of a term or word, within the syntactic context of a sentence (and previously Frege, Putnam, Bearling and many more, in the sense of identifying the semantic-meaning as an autonomous and objective characteristic, and not only with subjective attributes, which is the most common; “Meaning of the meaning” is a good reference for Putnam himself as an author).

For our part (B. Moreno, 2017), we have needed to now address the meaning, go through the concept of Cognitive Representations (CR), within the Three-Dimensional System of Cognition (TSC)[5], within disciplinary Psychophysics (with Stimuli -ES-, Sensations -SS-, and Cognitive Representations -CR-). Subsequently, we have addressed the Symbolic Capacity with characterization in the physiology of the nervous system (with correspondence in the components C1, C2 and C3, of the ACP -principal components analysis-, that has been carried out to establish the TSC), extending to the present study of *mathematical analysis of meaning.*

The signified-semantic framework is approached with a mathematical format from the perspective of W. Weaver, with stochastic probability processes, considerations at the entropic level of physical information, and successive semiotic evaluation, and in the last stage with P. Rocchi (2010 ), from the signal, both analog and digital, and with the attribute of *signified-semantic* without quantifying in its analysis.

---

[3] <information>: physical signal, its content is not considered. Linguistically it is the “lexical signifier”.
[4] Hinzen, Wolfram and Poeppel, David (2011). *Semantics between cognitive neuroscience and linguistic theory: Guest editor’s introduction*. Language and Cognitive Process. Psychology Press – Taylor & Francis Group.
[5] At www.ingit.es , in the Projects section, Neuroscience and Psychology section, thesis for a master's degree in research on cognition and behavior: 3- **Sensory-Cognitive Interaction in the Visual Domain**

Our contribution of quantification, first of the meaning and its theory, and then with the systematic application in the <Third social and economic sector>, carried out in the particular case of the Collective of Older People (explicitly in article 5.2, following this 5.1 ).
The aim is to provide them with maximum <cognitive accessibility> in this application; it is done seeking to go through a specific grid of verifications, which models the information with semantic signified, more widely than in the current multimedia, which being with a certain frequency "asignificant", is insufficient for the objective set forth here for supports in that Third Sector people in need of cognitive support.
The part that will follow the present study (5.1), which will be, as we have said, referenced as 5.2, will investigate the channels for the systematic extension of *<cognitive accessibility>* to the broad meaning. The objective is to serve cognitively disadvantaged people in the sector of activity referenced above, and in communication of all kinds afterwards, including the intellectually disabled, and in general in learning processes: in education, training and training, generalizing up to general disclosure and communication in the mass media in our knowledge society.

## 1- Mathematical theory of communication

In <*Recent contributions to the mathematical theory of communication*>, W. Weaver (1949), establishes a very psychological definition, *"(...) communication (...), as the set of procedures by which one mind can affect another*" , both orally and in writing, with music, with painting, theater, ballet and in general all human manifestations.

The levels of communication proposed by him are:
• Level A: as a technical problem of the transmission of symbols in communication.
• Level B: as a semantic problem of being received the symbols with the desired meaning.
• Level C: as a problem of effectiveness of the meaning to affect a desired behavior.

"A" is treated by the communication theory of C. Shannon.

We are going to focus on “B” in this study 5.1, and we study the identity or approximation of the meaning captured by the receiver, comparing with the meaning intended by the sender or source of information.

The problem of the "C" level of effectiveness is related to the semantic problem of A and B, on the one hand, and on the other it will be dealt with in the separate article that we refer to as 5.2, as it has been advanced.

W. Weaver determines that levels B and C are dependent on A, because it partially includes the two following levels. Even so, Shannon confirms that “*the semantic aspects of communication are irrelevant from the engineering point of view*”.
And in turn, Weaver clarifies "*but this does not mean that the engineering aspects are irrelevant from the semantic point of view*."

According to notes 2 and 3 on page 1 of our study, the term <information> in communication theory refers not so much to “what is said” as to “what could be said”, and adds , “**the <information> is the measure of the free choice of a message**”.

Between two possible messages to send, the choice of one underlies. The set of the two is the unit of information, and it can be thought that one of them is, but this is a flagrant error. It is the set of the two, the unit of measurement that governs/constitutes the <information>, which is the **Bit** (Bynary Digit)[6].

Therefore, in its simplest, irreducible base, it is a binary choice, a <<bynary digit>> (0-1), and that arithmetically represents the power expression: $2^1$ for a simple choice 0-1, or a bit ; $2^2$ for a choice between 4 possible options, or two bits; $2^3$ for a choice among 8 possible, or three bits; and so on, so the measure of a set of choices of,

- 2 cases, it is the base 2 logarithm, $\log_2 2 = 1$ bit
- 4 cases, it is the base 2 logarithm, $\log_2 4 = 2$ bit
- 8 cases, is the base 2 logarithm, $\log_2 8 = 3$ bit
- … etc

Thus, the logarithmic measure is totally natural, in terms of reason for being arithmetically typified, with an elective logic.

Let us take the 27 letters of the English alphabet as the possible choice between 27 messages, suppose each of the 27 letters/symbols to be equiprobable (which they are not exactly, but doing so for computational simplicity and a relative approximation), the information is of a worth,

- $\log_2 27 = 4.76$ bits

If to each letter/symbol, we attribute its probability of use in English $\mathbf{p_A}$ **,** $\mathbf{p_B}$ **,** $\mathbf{p_C}$ , … etc, whose sum is 1, we can establish that for each letter, taking its probability in "bits", it is the successive sum,

$$\mathbf{p_A \log_2 p_A + p_B \log_2 p_B + p_C \log_2 p_C} \ldots \text{etc}$$

(where $p_A$ is the probable portion of **A**, and $\mathbf{\log_2 p_A}$ is the bits of **A**, and so on)

Whose information in this case, with the accuracy of the probable occurrence of each letter (symbol), which is known for its statistical use, for example in the English language, the information that it actually provides is an average of 4 bits. This mean is less than the 4.76 bits with equiprobability, seen in the first formula and calculation.

This change is due to <redundancy>, which is essential for a message to be intelligible between noises that appear during transmission, delimited by repetition that disambiguates.

This quantity[7] is even lower if we do the calculation by rehearsing one-sentence sequences, giving us an even lower 2-bit bound on the actual informational content of the English alphabet.

If the maximum freedom of choice is 4.76, and the real one is 2, the ratio of the two measures is the relative freedom: 2/4.76 = 42%, so the redundancy would be 100 – 42 = 58% .

As a comparison in other languages, each symbol has a value informatively speaking:

- Morse code (dot-dash): **1 bit** (which is the $\log_2 2$)
- The decimal numbering code (0, 1, 2, 3,…etc): **3.32 bit** (which is the $\log_2 10$)

---

[6] W. Tukey usó el término Bit por primera vez. No confundir bits con bytes (combinación de normalmente 8 bits)

[7] J. Singh (1966). *Information theory of language and cybernetics*. Publishing Alliance 1972

Therefore, generalizing the formula of the symbols of the alphabet to the "n" symbols of any possible message among the collection of those messages, its generalized value is:

- $\sum_1^n$ $\mathbf{p_n \log_2 p_n}$ (in bits)

Shannon calls it <<**H**>>, and since $\mathbf{p_n}$ is a fraction of 1, the logarithm will be a negative quantity, for its improved expression in positive, with which the information H is expressed as:

- $\mathbf{H = - \sum_1^n \; p_n \log_2 p_n}$ [1]

Concepts such as noise, redundancy, use of parity bits to ensure the fidelity of the message, and other techniques and systems that accompany it in its transmission, in turn reduce the capacity of the channel to transmit symbols (messages) per unit of time, but they do it more safely. In this sense, we are interested in seeing the system in a block diagram, not going into details (which are vital for good transmission, but not necessary here). The communication in block diagram is as follows,

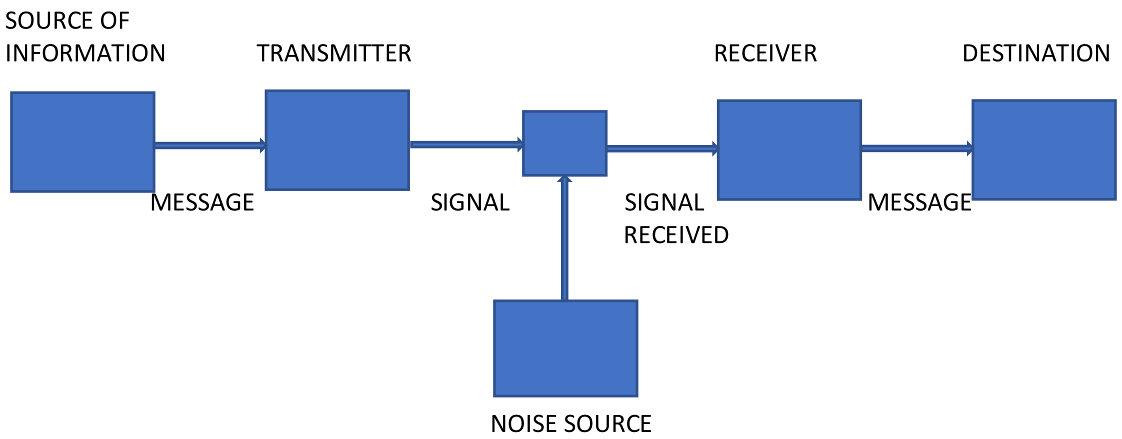



Figure 1

Continuing with Weaver's contribution, his suggestion (1st) regarding the previous block diagram, and referring to dealing with communication levels B and C below, this review consists of adding another box to the diagram that would be called "Semantic Receiver" , placed between the engineering Receiver that changes signals to messages, and the Destination. This semantic receiver performs a second decoding of the message.

Likewise, another block is introduced between the Information Source and the Transmitter, referencing it as Semantic Noise, leaving the name of Engineering Noise for the general noise source. We can see it in Figure 2.

A suggestion (2nd) by Weaver is that it is convenient to study the statistical structure of language, being in favor of using the theory of Markov chains, by adapting to one of the most important and difficult aspects of meaning, the influence of context. In our work here, we has taken

this point of view by considering the word as the first component of the language, and the phrase (as a string of words) as the context to provide more explicit meaning to each word.
As an "atom" of language, the word is the entity that syntactic rules cannot separate or reorganize.

The following suggestion (3rd), expresses that "*There is a vague feeling that the <information> and the meaning could be treated as a pair of canonically conjugated variables (...)*", an aspect that here, in this theory, is widely analyzed and shown in Section 3.

The next suggestion to be analyzed (4th), refers to the entropy introduced by Shannon in his theory as a parallel to information, considered as *"(...) is surely the most important fact*" at the "A" level. At information level B and C, its relevance will be made manifest again. It will be seen in article 5.2, continuation of this 5.1.

Finally (5th) states that *"(...) entropy not only speaks of the language of arithmetic, but also of the language of language*". Indeed, in this superposition, the metalanguage, opens the door to how this characteristic of child development arises in which, after the appearance of symbolic differentiation (Dual Representation 1, according to J. Delohache), a second metarepresentational perspective appears, radical for all information processes in a broad sense.
In the green curve of Annex 1, we observe the beginning and end at 36 and 48 months of Dual Representation 1 and 2, due to the ability of meta-representations and metalanguage.
It will also be discussed extensively in the commented article 5.2.

**Details of the five previous observations**

The 5 points adduced by Weaver are supported by the majority. All of them represent part of the proposals of this study, and that we are going to examine with these nuances and complements,

(1st) Extended block diagram with receiver and semantic noise, the communication block diagram, would look like this:

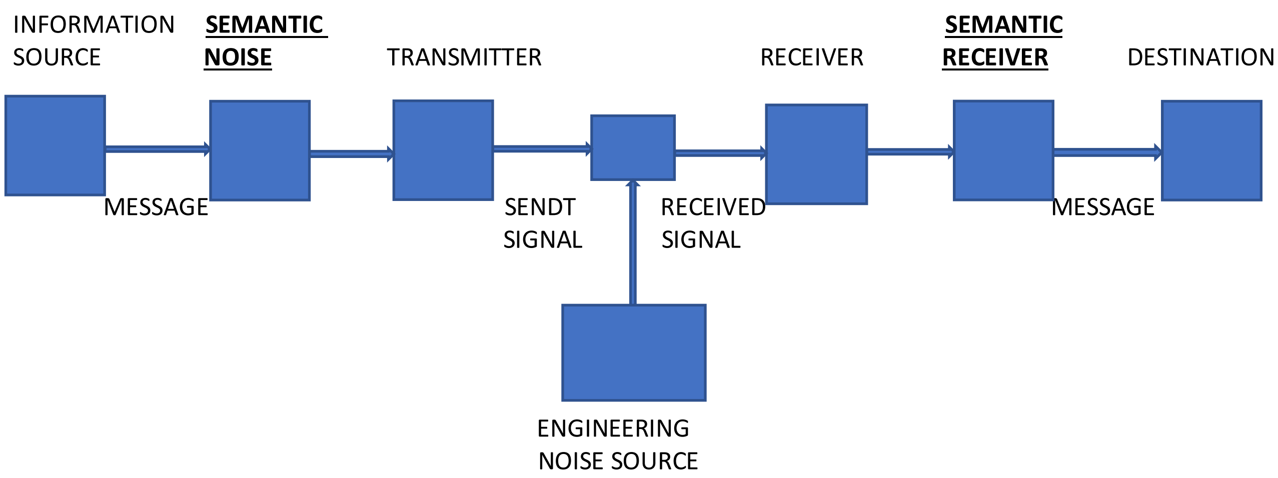



Figure 2

To specify the framework of reference of the model, we have chosen the emerging process in child development, of <verbal irruption> in infants-children from 18 months[8], with its learning curves that we will see in section 3.3, in the that appear both from lexical vocabulary (significant-lexical) and from vocabulary with broad meaning (meaning-semantic).
His analysis shows that if the first has a cost of learning level 1, the second has a resource cost of 3.5 for the same subject, for learning the meaning of his universe of vocabulary.

By introducing the boxes in the block diagram, of Semantic Noise and Semantic Receiver, these differential costs are being signified for the communication process that the "verbal irruption" supposes, and that occurs between the stimulating cultural environment and the mind of the user. child, formed mainly by his mother and family, and then by the entire nearby social network (in the nursery, the playground, etc.).

This happens in a continuous feedback of stimulation in the age window from birth to 48 months, and if this interaction does not occur[9], it can cause permanent injuries as in the case of Victor de Aveyron (raised as a "wild child ” in the forest without human contact) or that of children in orphanages in Poland during the Second World War, without any personal treatment.

(2nd) The phrase. Language with words (minimum unit of language, Bickerton[10], 2000) and with sentences (as combinations of these minimum units), are established as references for the evaluation of <verbal irruption> in infants-children.

The significant-lexical word, inserted in a sentence, requires from the baby-child a knowledge of the meaning-semantic of each term, much deeper in its meaning, and depending on the length of the sentence measured in number of words, and that create a meaningful context.

In this proposal by Weaver, reference is made to the analysis of sentences that contextualize the independent word, in a Markov chain. In our case, we directly value the sentence with that context potential of the words, and we focus on its purely semiotic and linguistic aspect, and linguistically a sentence, more than a chain, is a tree (syntagmatic structure).

In addition, there is the pure syntactic background, configured according to N. Chomsky's Universal Generative Grammar, with approximately 24 common functional rules of composition, existing in all languages, where the articulation of a sentence with sense or meaning is structured.

On the other hand, we start from the principle that in language as a cultural symbolic agent par excellence, in which "what is learned becomes instinct" according to the <Baldwin effect>, carrying behind it the entire biological evolutionary process[11], as well as anthropological that structures language, and that manifests itself in the development of each baby-child, immersed in culture, with probably a decisive influence on genetic and epigenetic aspects[12].

[8] At www.ingit.es , in development section, Neuroscience and Psychology section, Article 4, see growth curves .
[9] At www.ingit.es , under development section, Neuroscience and Psychology section, Article 3, pp. 9, 25, 26 and 28.
[10] Bickerton & Calvin (2000). Language ex Machina. Reconciling Darwin's and Chomsky's theories about the human brain. Editorial Gedisa. Page 37, 54 and 63
[11] Sampedro, J. (2002). Destroying Darwin. Ed. Critica s.l. 2002. – See note 10, Bickerton, p. 241
[12] See at www.ingit.es , under development section, Neuroscience and Psychology section, Article 3, pp. 10 and 26.

The set of factors establishes a narrow meaning in each *word*, and a contextual and broad one (Hinzen and Poeppel, 2011), with respect to the independent word, which is outside a sentence, clause, paragraph, text, etc.

(3rd) The arbitrary union of the signifier-lexical and the signified-semantic is an indicator of the independence of both symbolizations (F. Saussure, 1928).
Dimensions are the Cartesian mathematical construct, which allows the conjugation of variables, in this case that of <information> (in Shannon's terms) and signified in a two-dimensional, algebraic and geometric space, as independent variables in a two-dimensional plane.

P. Rocchi (2010)[13], exposes his approach to the semiotic triangle (Figure 3), as the optimal model, using within the linguistic model, the signifier (lexical) and the signified (semantic), on the one hand, and on the other the pole observer referent, which can be a person or an instrumental detector-actuator device that interacts with the significant and signified poles.
The latter allows, as long as it is evidenced by the object-event (OS) indicator, to have a magnitude that represents the OS in some way, as a hint, index, indicator, note, or as an icon, or simply arbitrarily symbolizing the OS . We will see it in greater detail in section 2 of this study.

(4th) But this "*Mathematical Theory of Signified*" would not reach the end of Weaver's proposals if he did not address entropy as a common thread in the world of physics with his second law of Thermodynamics.

From the polarities of Order-Chaos, theories of systems[14] have been developed, in some cases hierarchical, highly structured and ordered, as well as theories of chaos[15], the latter as a reference pole for entropy, in the same way that the systems in the other end they have it in the <information> (in the narrow sense).
Right in the middle[16] of these two polarities, the Theory of Adaptive Complexity arises, based on the emergent[17] adaptive self-organization, whose perhaps most systematic exponent is the Theory of Complex Networks[18].

A study center, such as the Santa Fe Institute, has had professionals from the fields of all these theories: *systems*, *chaos*, *complexity*, who have contributed essential concepts and approaches, and who pivot, in our opinion, on the physics of the second law of thermodynamics, that of entropy.
At the Santa Fe Institute, from Laughton to Kauffman[19] to Gell-Man; the physicist A. Solé, from the University of Barcelona, who was also at the institute for a second time, being on the other hand perhaps the most notable student of the physicist and museologist Jorge Wagensberg. Both in the Physics faculty of the UB, had an interesting relationship with Margaleff, the relevant ecobiologist of the UB who applied information theory to calculations of biological niches, who was followed by J. Terradas along the same lines and who currently continues F. Lloret, specialists in biodiversity.

[13] Rocchi, Paolo (2010). *Logic of Analog and Digital Machines*. P. 7, 32 and 38.
[14] Betalanffy, L.v. (1969). *Systems Theory*
[15] Gleick, J. (1987; 2011). Two texts: *Chaos* and *Information*
[16] Gell-Man, M. (1994). *The quark and the jaguar*. Tusquets Ed.. Pages. 77, 123
[17] Johnson, S.(2001). *Emerging systems*. Turner Economic Culture Fund
[18] Sole, R. (2009). *Complex networks*. From the genome to the Internet. Tusquets Ed., Metatemas.
[19] Kauffman, S. (2000). *Research*. Tusquets Ed., Metatemas.

Solé, with his work on Network Theory, with applications in linguistics, and complex systems in general, especially attracts the attention of this multidisciplinary study presented here.

Probably complex networks, and in particular the works of Erdos and Renyi, configure the space in which the theory of self-organized adaptive complexity converges, with a very strong component with very broad mathematical stochastics (such as that of Markov chains).
Of course, its relationship with linguistics is because it proliferates in superficial and deep structures ("s" and "p" structures), with phrase trees, with high levels of complexity.
Syntax is complex, but there is a reason for its complexity, and that is that thought is more complex than language, and we can only express it through limited methods, such as word-for-word language (in sentences, which are microcosms of words).
Grammar offers arguments for the refutation of the empiricist doctrine, that there is nothing in the intellect that has not previously passed through the senses.
In this discipline, the *Adaptive Self-Organized Emergence* on the Edge of Chaos between Information and Entropy, is where we find fractals (tree structures for language), the organization of many biological processes, and the tension of "coexistence" between the <entropy> and the <information>.

The excursus that we make here is very broad due to the importance, in our opinion, of the final objective that W. Weaver expresses at the end of his text, together with the relevance of entropy for the theme of signified, and that is the *language of language*.

(5th) The metalanguage is vital in the field of child development (Perner, 1991)[20], therefore, in addition to the treatment of <information>, entropy and meaning as Level B of communication in the Weaber classification, it is manifested the essential criterion of metalanguage.

Metalanguage allows reflection in thought (among others, Husserl deals with it in his Phenomenology), allowing recurrence both in reflective thought and in language speaking about itself.
It is the semiotic origin treated by U. Eco (1975) as "feasible to lie" (deinterlacing, entropy of the signifier-signified), which begins in the child's development at 36 months and ends at 48, like that of the dual representation 2 by J. Delohache. Lying presupposes des-interlacing, the entropy of the signifier-signified.
It also creates paradoxes, which are also resolved from the same metalanguage. From Tarsky's theorem it is done by untying Epimenides' "liar paradox", and much further by establishing the mathematical truth of <incompleteness> by Gödel's theorem. It is shown that any symbolic system is incomplete if it is consistent.

[20] Perner, Josep (1991). *Understanding the representational mind*. MIT

The common field of Tarsky's theorem and Gödel's, is in our opinion the metalanguage, whose basis is in the language modeled by the polarity of *signifiers - signifieds* (in terms of Weaver *[<information> - signified]*, central theme of this studio).

But without “wide semantic signified” and permanently open to expand, on any object-event (OS), that we point out, that we want to “narrate” in terms, there is no information conjugated with the signified. The metalanguage, by its self-reflection, finally allows us to have a theoretical and mathematical framework of reflexive animal and human consciousness, which can frame approaches for artificial intelligence (which today, in part, is captive of the effects referenced by the "Chinese room" J. Searle, Chalmers' "hard problem", Levine's "body-mind" separating gap, or/and Harnad's "grounding problem").
This part will be developed as has been commented, in Article 5.2, continuity of the present.

1- **Semiotic theory of information**, according to the <<*Logic of analog and digital machines*>> by P. Rocchi, 2010.

P. J. Denning, makes the prologue of this text with the following question by P. Rocchi "*What is information?*". Later in an article in "American Scientist", signed with Bell, they addressed the singularity of P. Rocchi's (2010) answer to that question, and it was: it is the information according to C. Shannon, and it is the signified[21], within the semiotic triangle “signifier - signified - referent”; and that this is associated with the significant information or <information> of Shannon; and it manifests itself linguistically and semiotically in the interpretation of the Referent (receiver or emitter <R>, who works with an *indicator or clue, or an icon, or/and a symbol* of the OS object-event represented, reason for the communication).

This relationship is expressed from the symbol as an ***artificial representation***.

The semiotic and linguistic appreciation allows us to see that "*The word is the quintessence of the symbol*" (S. Pinker, 1994)[22], we frame the nature of the symbol and the word, as corresponding.

The "symbol" is the term used by Turing to present the unit component of his universal processing machine (first computer proposal), also used by Shannon in MTC, in artificial intelligence by Pitts & McCulloch, in a long tradition of authors in disciplines when dealing with information (Gadner, Goodman, Deacon, Eco, etc.).
From the symbol, and from the word in particular, children have to tacitly assume that *they are not notes of the subjective behavior* of a certain person, but that they are shared in a shared way

[21] Denning & Bell (2012). *The information paradox*. American Scientist Nov.-Dec. 2012
[22] Pinker, S. (1994). *El instinto del lenguaje*. Alianza Editorial

bidirectional, and that to the speaker, it serves to transform meanings into sounds, and to the listener, the sounds into signifieds, both using the same code.

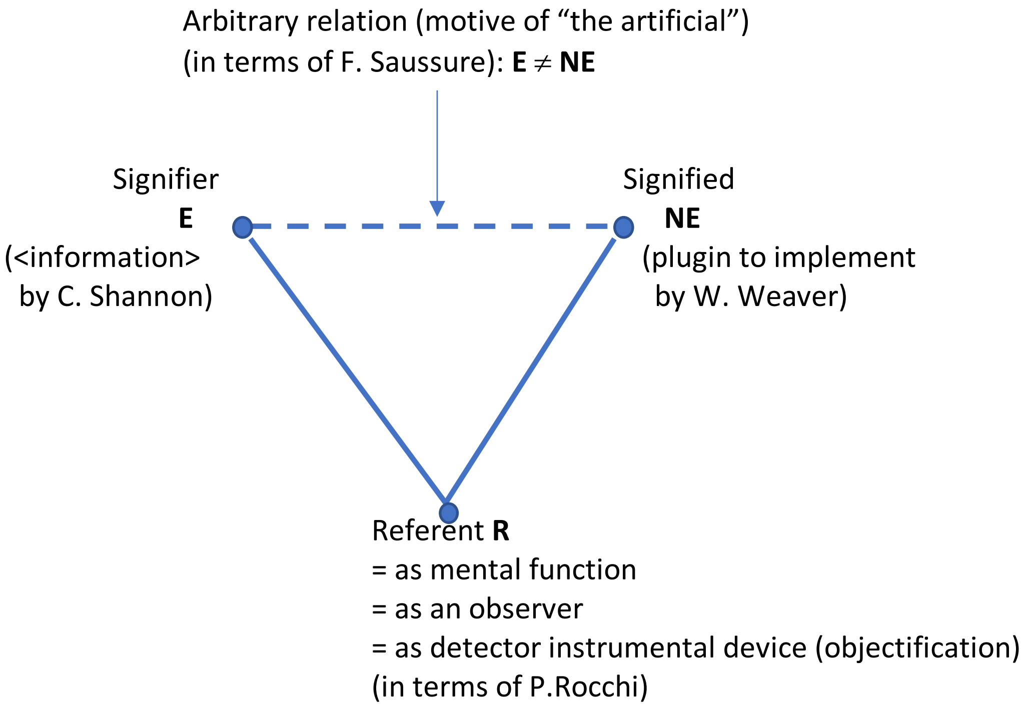
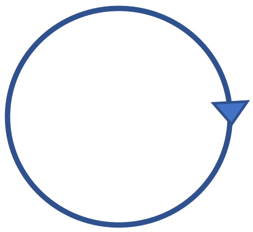



Figure 3

<u>The symbol converted into a circle from the previous triangle, to identify **E = NE** emerges as an object-representation of the **natural** OS</u>



**E** = **NE**
(in terms of P.Rocchi)

Figure 4

P. Rocchi, radically rigorous in the examination of semantics, exposes making *"(...) an independent <u>attempt</u> to dissect information technologies on the basis of <u>physical properties and semantic concepts</u> as well*" (p.40). The underlining is ours.
And it goes on, *"Subjective factors interfere with semantics"*, not seeming obvious by this factor, to leave semantics aside. He then cites three fields of knowledge that have been fruitfully addressed despite the implicit subjectivity they have (such as semantics):

1) <u>The probability p(A)</u>, in the Bayesian methods of the subjectivist school with p(A), which expresses the degree of belief of an individual about the occurrence of A.
2) <u>The reference systems in the movement of kinematic physics</u>, in which the speed of a body depends on the observer, in the sense that two observers can assign very different speed values (e.g. movement of a train on the parallel track with respect to another train, who is moving, him or us?).

3) Medicine, as a personalized discipline; for example, two patients treated with the same protocol sometimes differ drastically, when one dies, and the other survives (also e.g. due to subjective tolerance to the pharmacopoeia and to all medical remedies; and for this reason alone, it did not take experts to abandon the study of medicine).

Continuing: *"The aforementioned scientific challenges are not intellectual curiosities or tiny trajectories belonging to the scientific mosaic."*
*"Scientists cannot destroy the subjectivism of natural phenomena, however they do not give up and work hard to improve knowledge of phenomena and delimit the effects of arbitrariness."* We could not agree more.
*"I see no reason to avoid parallel efforts in information science, and this book represents an attempt to show how semantics can be discussed on the basis of objective problems (...)".*

The digital can reconstruct the analog, but the reverse is not true, because it is more blurred, less “sharp” (P. Rocchi's principle of sharpness). But it must be said that analog is what we handle, as organisms and humans, more usually.
For example, with language, the most relevant symbolic process with which we deal, however, has a fractional (if not digital) component, notable: letters, words, phrases, paragraphs, texts, etc. But this does not prevent us from having to use the "mid-air" with an acoustic signal, intrinsically analog, to hear and emit the words.

Rocchi establishes an even more progressive succession in levels from 1 to 5 for the informational structure:

1) Bit's
2) Binary words
3) Common words
4) Texts (and images and sounds), and
5) Multimedia and hypermedia information

The migration from the most fractional to the most holistic information also follows the trajectory from the most digital to the most analogical, and from the most significant to what has the most signified (for example, from level 3 to 4, to go from the words to texts with sentences, on the referenced scale). This progression, Rocchi calls it <<*assembly*>>, and has a great conceptual similarity with what we have here called <intertwining>, to characterize the relationship of the two different dimensions, but intimately related, of signifier and signified, which will be seen in Section 3.

It is not a term of excessive conceptual construction. He says “*a web page that offers texts, musical pieces, maps, diagrams, movies and interactive tools constitutes a good example of hypermedia communication*”.

### **2.1 <The signified>, as a progressive, incremental and unlimitedly feasible term *to be increasing in meaning***

Technology is represented by the machines that form it.
Machines manipulate <information> and meaning in ever-expandable ways.

They allow us to understand a step further in organs such as the brain and its mental function, and coding; Likewise, also possibly in machines such as computers and their processes, and their similarity with the thought process, with Miller's statement about recoding: <<*it seems to me that it is the vital sap of mental processes*>> (Gleick, 2011, p.364). Let's look at various operating devices in computing,

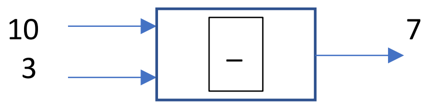
21.1 Converters

211.1 One key: makes the mechanical-electrical analog conversion.
211.2 A led: converts electrical signifiers into light.
211.3 A printer: converts electricity -> print
211.4 A screen: electrical bits -> pixels
211.5 A loudspeaker: electric waves -> sounds
211.6 Un CCD (charge coupled device camera): light signal -> electrical
211.7 An imam
211.8 An electric motor
etc.

In all of them, the conversion of the signifiers from one to another takes place ($\mathbf{E}_{input} \neq \mathbf{E}_{output}$), keeping the signifieds unchanged ($\mathbf{NE}_{input} = \mathbf{NE}_{output}$); see in 2.2.

21.2 Data processors (create new sets of <information> terms, phrases, contexts in texts, languages, etc.).

212.1 Numerical. Performing numerical calculations




212.2 Verbal: An example on the Internet with the “Search” function, to find “a mobile phone”, we go to -> Websites with different telephony offers.

212.3 Visuals:
- sequences of successive images, such as videos.

212.4 Simulations:
- of airplanes in flight, in chemistry, physics, etc.
- financial statements
- news
- management of sounds, texts, letters, images

212.5 In the real:
- precise analysis of the meaning in the "practical experience", with numbers that represent it (with a label)
- Dollars
- euros
- liters
- products... etc.

212.6 Creations:
- the machine does not invent, but has creativity, providing original ideas, for example attributing or deducing meaning as from 212.1 to 212.5

212.7 Conversions into new data and into reality:
- "*How can a computer system that is unconscious and does not know the meaning of its data, transmit novel information*?"

⇨ By instructions applied to the databases, from the memories of these data and establishing relationships and correlations (for example between the space traveled, and the time spent traveling it, the speed ratio gives us, new information not obvious from the data starting space-time).

212.8 Software developers
- They can: handle instructions and databases by hand, not in a fixed way, but creatively.
- cannot: implement an "effective software" application in a simple way, with a rigid metrology.

**2.2 In a computer, E (significants) and NE (signifieds) are handled, using different units of operation in different configurations**

- Working together **E** and **NE** (as we already saw as an introduction, in section 21.1)

(1) Varying only the signifiers: ($\mathbf{E}_{input} \neq \mathbf{E}_{output}$)
and keeping the signifieds unchanged ($\mathbf{NE}_{input} = \mathbf{NE}_{output}$)

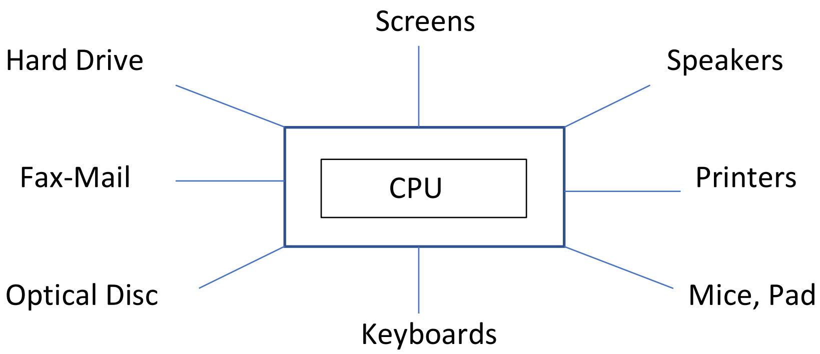
Being **(1)**, therefore a **non-informative** conversion (only the signifiers vary)

(2) Only the signifieds vary: ($\mathbf{NE}_{input} \neq N\mathbf{E}_{output}$)
and keeping the signifiers unchanged ($\mathbf{E}_{input} = \mathbf{E}_{output}$)

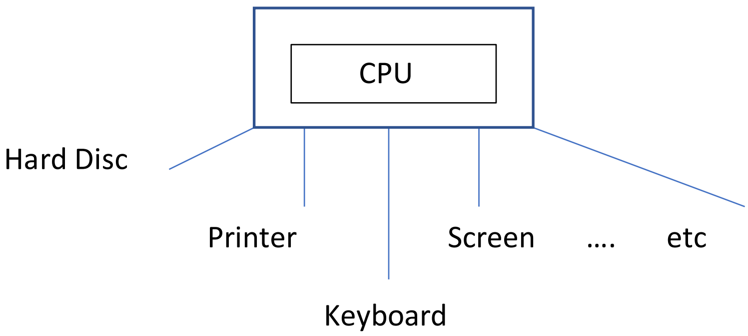
Being **(2)** novel, **informative**.

Due to the structures or architectures of the technological devices linked together, they define the possibilities of <digital convergence> in different formats, as well as with a variety of **relative informational capacity:**

22.1 Star (CPU, Control Process Unit): radial model that is clearly facilitated by the USB (Universal Serial Bus) inputs and the network of networks, the Internet,




22.2 In tree: commanded hierarchical model; such as mobile phones, hi-fi i-pods, game controllers, satellite navigation,




22.3 In chain (IPO model; Input, Process, Output):

It is frequently seen in agricultural production, water, mining..., but also in number processing, images, sounds...


It is a frequent mistake, in the Computer Science discipline, to prepare technical manuals as "cookbooks", when they must communicate the <theoretical computer science principles>, established today preferably in the digital paradigm, both in its bases and in its objectives. , not with a recipe book, as is done too often.

The CPU and its structuring of peripherals is the key, since its processing, its combinatorial capacity, articulates the processed information in a different way, as does the speed formula when dividing space by time (section 212.7)

22.4 Scaling levels 1 to 5 (page 12): from the simplest signifiers, such as Bits or binary words (at level 1), to indicators with sensory perceptible cues, such as multimedia/hypermedia (at level 5); which is not simple, much less obvious.

At this level 5, the audiovisual is significant: from the indicator-indication shown in a correlated way with a text, to the systematic labels, or good correlations with narrative texts, etc.

The pure "conversion" of code (according to formula **(1**)) does not contribute semantic-signified; when "instructions" intervene on a database (according to formula **(2)**), we can have an output of the set that provides new meaning, forming semantic-signified: by combinatorics, by aggregation, by syntax, etc.

In conclusion, and also as a summary here on "*Logic of Analog and Digital Machines*", P. Rocchi's proposal is developed around information in ICT (Information and Communication Technologies), with the concepts of <Paradigm Digital> preferably, but also assuming the essential part of the <Analog Paradigm>.
Treated both for all kinds of devices, instruments and machines. Through them, it is established in human processing, first from the point of view of the brain as the central nervous system, and previously from the peripheral nervous system, which coexists in interaction and interface with the peripherals of computer machines (although internally to the animal and human organism, action potentials, brain "spikes" are clearly "digital", but without natural interaction with the outside - except through instruments such as EEG, MEG, functional MRI, etc.).

For this part, taking in our case, the human symbolic system, the most important, language and its semiotic theory, to develop the theory of information, becomes central, because it incorporates the signified factor as essential, despite brushing with it elements of subjectivity that could distort the scientific approach, something that P. Rocchi satisfactorily avoids and argues, in our opinion (in the header introduction of this Section 2).

His approach to information puts us in a position to address the next section, on the other hand, in order to move towards a greater typification and quantification of meaning, including criteria for an expanded information theory (with semantic-signified and its meaning "broad").

## **3. MTS - Mathematical theory of signified** (from the related <STC> theory – B. Moreno, 2017).

The framework of the information theory of P. Rocchi with the starting linguistic and semiotic approach, establishes the bases to advance in what were approaches in the "*Mathematical Theory of Communication*", basically in the writing that accompanies that of C. Shannon, by W. Weaver in 1949.

The approaches of W. Weaber (see Section 1 page 3), are established as a <contribution> to Shannon's theory, and are very relevant regarding the incorporation of meaning for communication theory. Here we will deal with the first three, (1st) Changes in the <block diagram> (Figure 2), of Shannon's communication (Figure 1), (2nd) Statistical structure to investigate the language, and (3rd) Variables canonically conjugates of "information" (composed of signifier-lexical and signified-semantic, in our terms).

The approaches of the proposals (4th) and (5th) will be developed in an article in continuity with this one, referenced as 5.2.

3.1 Changes in the block diagram of C. Shannon (1948)

31.1 The Semantic Receiver

The reception of the message from the point of view of the <semantic signified>, supposes several stuff,

a) That the message carries encoding with a semantic dimension
a1) If the message is a word, the semantic code could be a phrase that contains it with the meaning to be transmitted (of different options, as it can give the meaning: a dictionary, giving an implicit definition in a phrase that contains that word )
a2) Following a word (called in linguistics “holophrase”), the semantic code could be an image or a sound or both, which express the meaning of the word in transmission.

b) If the above were indicators or indications of the <object-event> OS that names the word, a different situation would be:
b1) Name a synonym of that word.
b2) Attach an icon or ideogram (similar to a Chinese letter, or emoticon)
b3) The function of attaching other symbols, such as: acronyms, with acronyms (for example UK, from United Kingdom), or that term in another language (translating it, United Kingdom).

c) Also if it refers to a more or less technological device: a dictionary in the form of an encyclopedia, a voice assistant, a definition on Wikipedia, etc.

Based on the previous conjugate associations, the concept of “interlacing” (<assembly> in Rocchi's concept and next), is fundamental here; it is the interrelationship between the “lexical signifier” and the “semantic signified”. Many variants appear to establish it. Two options from here treated, additional and quantified, such as:

- If we encrypt it in cost of learning (of analog characteristics)[23] .
- If we value it in terms of Shannon's <information> (of digital characteristics), as we have referenced it with an additional block of the <Semantic Receiver> (in Figure 2, details will be seen in point 3.3).

In both cases, W. Weaver's Semantic Receiver has two contexts, one analog and other digital.

After studying P. Rocchi (2010), in Section 2 above, we opted for the <digital paradigm>, if what we want to do is "communicate", but we would do it with the <analog paradigm> (sensory input) , if we intend that the receiver <learn> directly, for which it is even possible, that in an overall vision, it was done with both, communicating and learning, within its functions.
The Semantic Receiver would be left with these attributes, quite complete.

[23] On the web www.ingit.es , section Under development, Neuroscience and Psychology section, Article nº4

31.2 Source of Semantic Noise

In the block diagram of Figure 2, it appears in highlighted boxes, and will affect the semantic part exclusively, and apparently, only in the communication process.

It is important to note that it will affect the associated semantic package (phrase, image, sound, synonym, etc.), which will initially have a noise level; Shannon's principles and theory regarding this exclusive package, respond similarly to "Engineering Noise".
But the "intertwining" speaks of the rules of interaction of the signifying part with the signified part.

Most of the common words have more than one meaning, and there are few meanings that have more than one word, that is, homonyms abound, while synonyms are scarce.
It is not known exactly why languages are so lax with words, and so generous with meanings. It is possible to be a principle of minimal physical effort.
All this configures a potential semantic noise.

Can these rules be affected by a noise other than Shannon's?
Can an instruction of some computer program be affected by some other noise?
Can the GGU (Chomsky Universal Generative Grammar), in a human receiver, not be infallible for "interlacing", despite being basically biological? Or even simpler, can the syntax of a sentence be affected by the entanglement? noise?

Darwin said of syntax that it is the fruit of <an organ of extreme perfection and complexity>.
The organization of grammar should be in the developing child's mind, forming part of the language learning mechanism, which allows the child to make sense of the noises he hears his parents make.
The complexity of the mind is not a consequence of how learning is processed, rather, learning is a consequence of the complexity of the mind (Pinker, 1994).
And we are not even talking about the "intertwining" of the word-message with an image, a sound, etc., even more complicated to typify, although possible to carry out.

The answers require research that will not be addressed here.

## 3.2 The statistical and algebraic structure of language, both of the Signifiers, and Signifieds.

### 32.1 Concept like that of Markov chains and the like.

321.1 The context of a sentence, analyzed only by the combinatorial process (with the Markov chain property), is insufficient with respect to the human sentence-forming system (much more selective).
However, for speech recognition, the chain of probabilities can help for assignment/confirmation of feasible sequences. In fact, everything advanced in voice recognition technology uses mathematical resources derived from the probabilities of using terms.

321.2 The chain stochastic process.
It requires the assessment of the probability that one word follows another. Another variable of the process is the probability of appearance according to the length of the word, and many more studied in linguistics, such as the application of Zipf's law.

32.2 Syntactic trees (S. Pinker, Bickerton etc).

322.1 Phrases and relations
The phrase is a group of words that work as a unit in a sentence and that usually has a coherent meaning. Generative grammar uses a set of rules that determine the form and meaning of the words and phrases of a particular language, as spoken in a community.

322.2 GGU Rules (Universal Generative Grammar) according to N. Chomsky.
It is the basic, common design to the grammars of all human languages. His discipline is generative linguistics. Underlies a single machinery in charge (physically located in the brain in the <left perisylvian region>, including Broca's and Wernicke's areas, whose injury produces aphasia), and responsible for manipulating symbols, which without exception, cements all the languages of the world .
It should not be confused with the phrase structure surface -s, and deep -p, which are the grammatical categories of rules that form a sentence, usually in tree form as well.

322.3 Fractal structures.
Of recursion as an essential factor (abounded by Hofstafter with texts such as "Gödel, Echer, Bach", "A strange loop" and "Analogy"). Recursion is used to apply repeatedly to a different group of elements, such as words, that can be created and analyzed in any size. For example, order words by “alphabetical order” following the sequence of letters of the alphabet.
In phrase structures (syntagmatic), words are grouped into phrases, and these are grouped into *larger phrases*. They are *normally represented by a tree-shaped diagram, which is inherently fractal*.

322.4 The structured specificity of mathematical spaces: algebraic space ($\xi_a$), metric space ($\xi_m$), and vector-geometric space ($\xi_v$), are successive expansions, mathematically equivalent in some cases (the arithmetic of numbers with the geometry) and in other emerging ones (such as the effects of new dimensions, for example of the imaginary numbers of the real ones).

3224.1 Algebraic space ($\xi_a$)
It configures the “identity” or “equality”; the equivalence, the analogy, the <sharpness> of the terms.

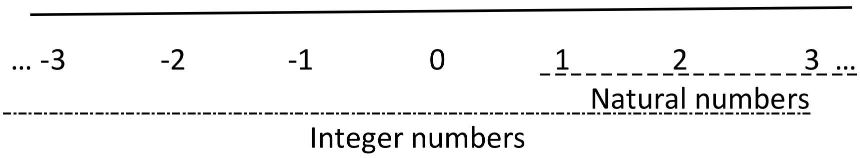
$A = A$

$A \neq B$

Entity, $E \neq E^*$ (contiguous differentiated; P. Rocchi, 2010)

NE (not E; P. Rocchi, 2010)

3224.2 Metric space ($\xi_m$)
Configure unit definition; comparison and quantification.
The natural numbers, the integers and the associated coordinate are reflected.




... -3 -2 -1 0 1 2 3 ...

Figure 5

3224.3 Vector-geometric space ($\xi_v$)
It is the space of the initial Cartesian plane, and the geometric space as an application (all possible two-dimensional geometric figures can be represented).
Real, imaginary and complex numbers can be represented in two orthogonal and associated coordinates (ordinate and abscissa).

Imaginary number (i)

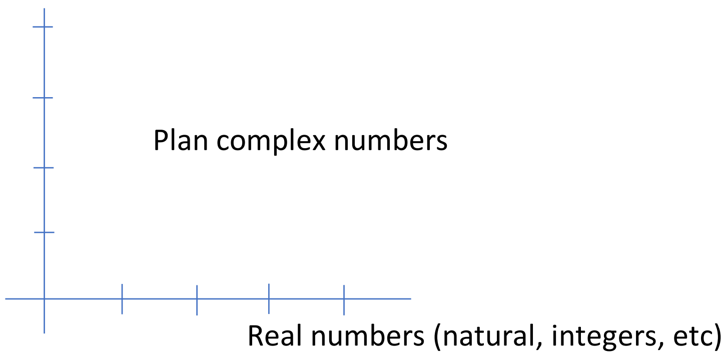



Figure 6

The Cartesian model is probably the best mathematical model to "conjugate" (W. Weaver's signified), the <signified> and the <signifier> from the P. Rocchi model.

Taking the semiotic triangle seen in Section 2, Figure 3, and projecting it onto the Cartesian coordinates in the vector-geometric space ($\xi_v$) of the complex numbers, with the NE and E poles, in the plane.

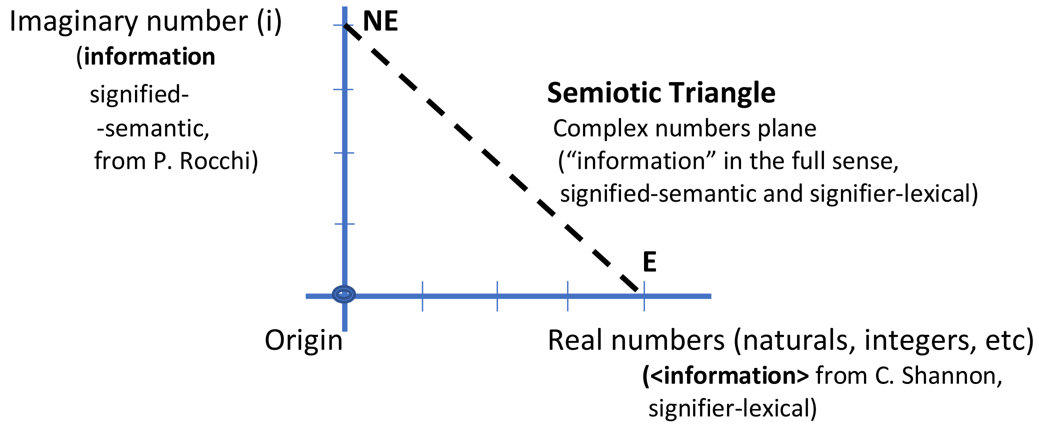



Figure 7

In the Origin of coordinates, we have the "zero representation", that is, in one sense is the Object-Event (OS), reason for the Reference (R), and in another sense the observer, the device-instrument-detector of the OS , in the objective sense (a voltmeter, a thermostat, etc., indicator of the present object-event), which is delimitable without being the OS itself, although it is encoded as a represented entity.

We agree that at the origin the equivalence R ≡ OS occurs, beginning in it the symbolic representations E and NE, with the character of a codified representational entity (objectively or agreed), and a different accent such as: [*indexes-indicators*], [ *icons*] or [*symbols*] of entities E and NE (Bearling 2009, Rocchi 2010, and others).

See regression lines in graph Fig. 8

3.3 <u>Shannon's <information> and signified, "canonically conjugated" (in the proposal of W. Weaver, 1949)</u>

Our objective here, following that proposal, and starting from the learning graph to typify the "linguistic irruption" in infants-children (B. Moreno, 2020; here in note 24 and graph in Annex 1), is to adopt in the code of natural language what links "signifier" wiht "signified", and as we have seen reflected in the algebraic space $\xi_a$ , metric $\xi_m$ , and vectorial $\xi_v$. But let's see how it is linguistically structured in a term equivalence table:

| **LEVEL** | **AUTHOR** | **FRAMEWORK** | |
|---|---|---|---|
| 1 | Saussure<br>Peirce<br>Morris | **SIGNIFIER** | **SIGNIFIED** |
| 2 | P. Rocchi (2010) | Signifier Semiotic **<E>** | Signified Semiotic **<NE>** |
| 3 | C. Shannon (1948)<br>W. Weaber (1949) | <Information> | Signified |
| 4 | Propousal 1<br>(ours, with desanbiguating purpuse) | **Signifier-lexical** | **Signified-semantic** |
| 5 | Propousal 2 | Isolated word | Word in phrase; contextualized |
| 6 | Hinzen – Poeppel (2011 | Signifier with <**narrow signifier**> | Signified in a sentence, <**broad signified**> |

| | | | |
|---|---|---|---|
| | ( in a similar sense, Putnam (1985) Bloom (1985), etcétera) | | |
| 7 | Propousal 3: **"verbal irruption"**, in baby-child development (see Annex 1) | **Number of vocabulary words** handled without context, isolated. | **Number of words in a sentence**, articulated in context. |

Table 1

Level 7 with proposal 3, is the basis that we will take for the quantified evaluation of the information with signified. They are applied as indicators depending on their origin, the isolated word and the contextualized word in a sentence.

The "verbal breakthrough" in children at 18 months, emerges with narrow meaning with the approximately 75 vocabulary words, learned and pronounced. They can also be thought of as single word phrases or "holophrases", which have a very low broad meaning, in fact referred to as <narrow>. In this context, morphemes are the minimum units endowed with meaning, into which words can be broken down (in- /micro- /onda- /bili- /-dad), giving meanings that can be composed.

**Summary of tabulated data on the development of a baby-boy in months of age:**
(ln: natural logarithm, to analyze regression)

**SIGNIFIERS**

| Word number | ln – Word number | Age (months) |
|---|---|---|
| | | |
| 0 | 0 | 13 |
| **75** | 4,32 | **18** |
| 375 | 5,93 | 24 |
| 1050 | 6,96 | 36 |

Table 2

**SIGNIFIEDS**

| Word number | ln – Word number | Age (months) |
|---|---|---|
| | | |
| 0 | 0 | 15 |
| 1 | 0 | 18 |
| 2 | 0,69 | 21 |
| 11 | 2,40 | 36 |
| 18 | 2,89 | 48 |

Tabla 3

33.1 <Intertwining>, or interaction relationship between signifier-lexical ST (E), and signified-semantic SD (NE).

We are going to define next, the <information> in terms of Shannon, and both of <E>, and of <NE>. In the linguistic field, E and NE, are the number of single words and the number of words combined in a sentence, whose magnitude delimits the number of words that compose it.

We call these two magnitudes HE and HNE; respectively, information provided by the number of signifier symbols, for a developing child coded in months, and the number of signified symbols of the number of combinable words in a sentence at successive ages.

**Regression curves** between **E** and **NE** show their mathematical independence as variables

| **Age** (months) | **ln ST** (signifier) | **ln SD** (signified) |
|---|---|---|
| 13 | 0 | |
| 15 | | 0 |
| 18 | 4,32 | 0 |
| 21 | | 0,69 |
| 24 | 5,93 | |
| 36 | 6,96 | 2,4 |
| 48 | | 2,89 |
| | | |

Table 4

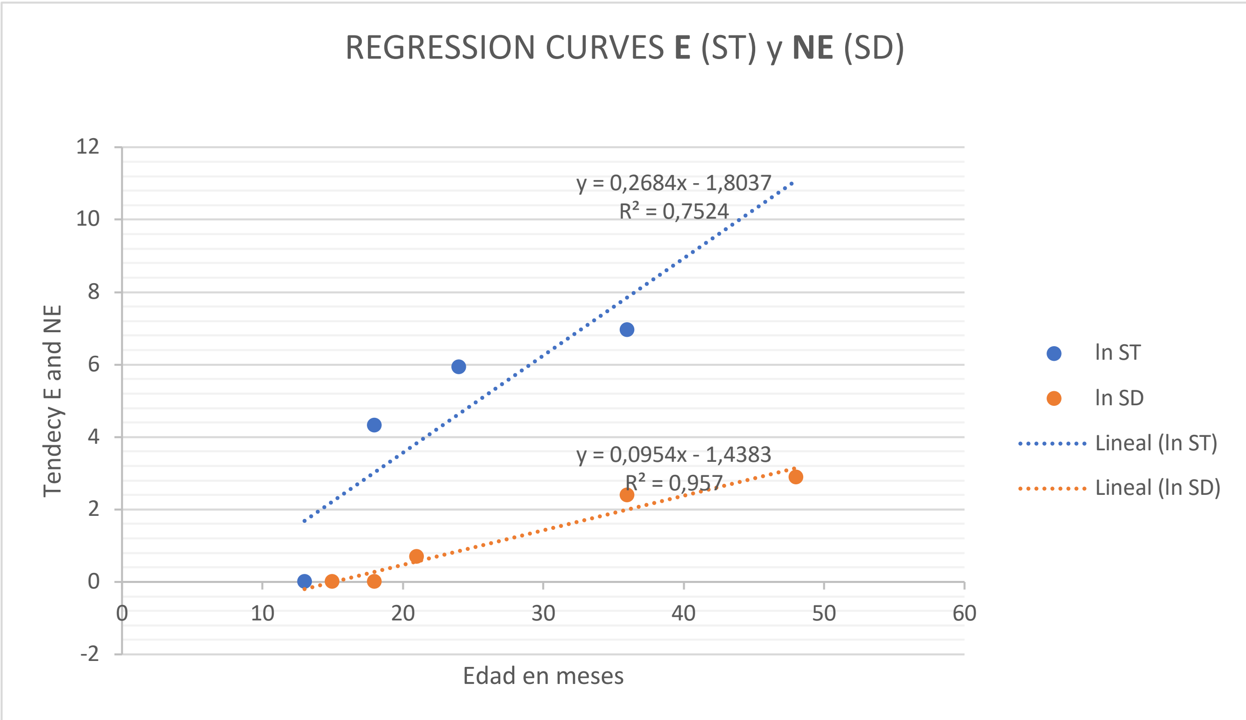



Figure 8

At approximately 24 months, children come to have about 375 words, and with sentences made up of 3 words in each sentence in which their "composition" begins (trained monkeys, they link a maximum of tens of 2 terms or comprehension words in each sentence, many of them gestural).

NE–E entanglement, like HE and HNE information, for an expanded information theory with the <broad signified>,

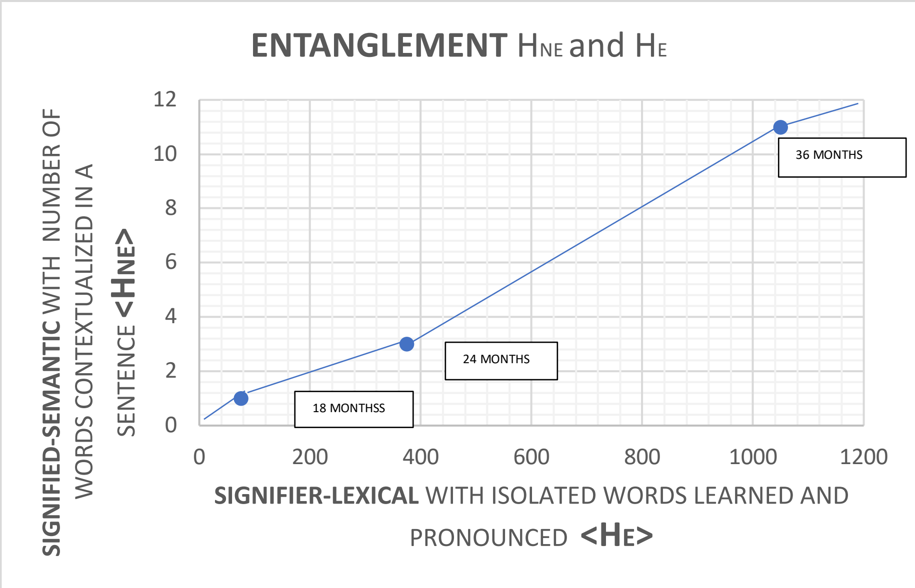



Figure 9

Each word is a symbol that goes in the communication message in Shannon's model, therefore Shannon's <information> for HE : 375 words, and simplifying considering all equiprobable words,

$$\mathbf{H_E} = \log_2 375 = \underline{8,55 \text{ bit}}$$

In the case of signified, its informative potential is given by a maximum level of HNE : 3 words, but which are combined from the 375 significant words E.

We therefore establish at a combinatorial level, how many sentences can be created without restrictions with:

n = 375 words to combine
r = 3 words combined in each sentence

The combinations are then:

$$\mathbf{C_n^r} = \binom{n}{r} = n! / r! (n - r)! \approx 8.859.500 \text{ sentences of 3 words}$$

Therefore, the information with contextualized signified in 3-word sentences, being simplified equiprobable:

$$\mathbf{H_{NE}} = \log_2 8.859.500 = \underline{23,1 \text{ bits}}$$

If we go to our algebraic space $\xi_a$ , metric $\xi_m$ , and vector-geometric $\xi_v$ , to represent it, we define the “conjugate information” or “interlaced”, as **HH**.

For a given object-event OSn, for each of the “n” words, which in this case is the 375-word significant vocabulary, we have a contextualized level in 3-word phrases.

The possible signified of these 3 words in a child with a developmental level of 24 months,

$$\mathbf{HH}^{24}_{375/3} = \mathbf{H}^{24}_{375} + i\,\mathbf{H}^{24}_{375/3}$$ in the adopted plane of complex numbers

$$\mathbf{HH}^{24}_{375/3} = \mathbf{8{,}55} + i\,\mathbf{23{,}1}$$

33.2 We extend entanglement as a <conjugate system> (W. Weaver), to the same developmental phenomenon in 36-month-old infants.

The non-linear fraction, going from 375 words at 24 months, to 1050 words of vocabulary (significant-lexical) at 36 months, considered with a different number of words, and also establishing the meaning-semantic terms, in which sentences of 11 words are created in those 36 months.

Following the guidelines of 33.1 but with a different number of words, and also considering its non-linear,

**n** = 1050 words to combine
**r** = 11 words in each sentence

We can compute Shannon's <information>, which is:

**HE** = log2 1050 = <u>10,04 bits</u>

And in an information with broad signified, the combinations are then:

$$\mathbf{C}_n^{\,r} = \binom{n}{r} = n!\,/\,r!\,(n-r)! \approx 4{,}223 \cdot 10^{28}$$ frases de 11 palabras

The information with contextualized signified in sentences of 11 words for an equiprobable simplified format (as in the "protolanguage", in the descriptions of Bickerton & Calvin, *Lingua ex machina*, p.53) for each word,

$$\mathbf{H}_{\mathbf{NE}} = \log_2 4{,}223 \cdot 10^{28} \approx \underline{95 \text{ bits}}$$

The **conjugated or intertwined** information **HH** in the mathematical space ($\xi_a$ , $\xi_m$ , $\xi_v$), remains,

$$\mathbf{HH}^{36}_{1050/11} = \mathbf{H}^{36}_{1050} + i\,\mathbf{H}^{36}_{1050/11}$$ in the plane of complex numbers

$$\mathbf{HH}^{36}_{1050/11} = \mathbf{10} + i\,\mathbf{95}$$

Whose generalization is obviously the same, and which is reflected in the formulas:

Generalizing: : $\mathbf{HH}^{R}_{OS_n} = \mathbf{H}^{R}_{E} + \mathbf{i}\,\mathbf{H}^{R}_{NE}$ [4]

(equation in the plane of complex numbers)

Being,
**R:** Reference, for an Observer, Device-instrumental-detector-actuator,
according to index-indicator-indicator / icon / symbol (Rocchi, 2010), of a given OS "n"
**OSn**: Object-event "n", origin of the symbolic representations "n"
**E**: Linguistically symbolic entity <signifier> (Rocchi, 2010), of the OS
**NE**: Linguistically symbolic entity <signified> (Rocchi, 2010), from OS
**i** : imaginary unit of the complex plane, in the vector-geometric space □v.

The formula displayed based on the probabilities of use of each representation symbolic "n",

$$\mathbf{HH}^{R}_{OS_n} = -\sum_{1}^{nR} p_E \log_2 p_E - i \sum_{1}^{nR} p_{nNE} \log_2 p_{nNE} \quad [5]$$

33.3 Appreciations for the specific content of the message.

With what was discussed in 33.1 and 33.2, and we have evaluated the information channeled during the verbal irruption of language at 24 and 36 months.

Taking at 24 months, $\mathbf{HH}^{24}_{375/3} = \mathbf{8,55} + i\ \mathbf{23,1}$ (in bits)

We can adopt a specific content of the 3-word sentence, for example, with the following sentence:

**"Dad is happy"**

Said by the 24-month-old baby-child to her mother, and that once formulated (to have the explicit meaning that is wanted to be transferred to the mother), it can be quantified.
To do this, the lexical phrase with 8.55 bits of information from Shannon and 23.1 bits of information with meaning is sent through the communication process, represented by the "block diagram" (Figure 2).

Everything happens within a syntactic code formed by: name, verb and complement.
It is shipped with a permutation capacity of 3 specific words, 3! = 6
The information associated with these 6 combinations represents an informational value of **2.6** bits, which compared to 8.55 represents **30%** occupancy of the potential communication of the channel in the lexical fraction. Semantics contains other factors of possible definition, let's see some possibilities, balanced from the maximum 23.1 bits of the possible semantic beam.

It can carry the following lexical associations:

- "**Dad**": "*male parent*", in which neither "parent" nor "male" belong to the verbal universe of the 375 words that in this stage of development it has at 24 months, and that the child handles. For the

composition of the sentence, several words have to be chosen from the dictionary of the 375, or direct attributions that associate "dad" with that person who is at home.
- "**is**", present of the verb "*to be*", factual state of that person.
- “**happy**”, “*emotional state of satisfaction*”, possibly also by direct attribution.

o Another reception option is to associate one or more specific sensory modes:
- For example, photo of the father smiling
- For example, audio recording of the father laughing
- A hybrid option of abstract (linguistic) – concrete (sensory) format: mixing verbal and sensory modes.

Finally, a balance will be made of the <specific message>, based on its lexical universe (of 375 words at the level of development with 24 months) and the terms of the dictionaries that we attach (lexicons with more words; or associating sensory messages in ways visual, auditory, etc.)

If for each word of the message, we explain it by at least another 2, and at most up to the remaining 375, the result is minus 3 of the explicit phrase, we are talking about 375 – 3 = 372 known terms to choose for the two additional terms .

The "little world"[24] of this dictionary is quite restricted, but it is the one that the child now has. However, it must be remembered that the memory "accumulation" or "clusters" are also used, which in the form of engrams have sensory modes inscribed, not lexical, which have their own dictionary (visual, auditory, etc.) and which can be hybridized with that of the lexicon when they are evoked.
The latter do not go through the oral channel that the child is using with his mother, but not being present in the communication, if they are in the interpretation of information that the child makes when launching the spoken observation to his mother, basically fulfilling the functions of semantic receptor.

What is not lexical or semantic does not enter into this assessment, limited to linguistic terms.

**CONCLUSION**

Equations [4] and [5] describe and quantify, without entering into considerations of redundancy, security or complementary aspects of interpretation, the significant-lexical information and that of meaning-semantics, which allow determining and completing the informational value of the meaning. that concerns us, in the second term of these equations.

This value is objectively, and taking the "verbal breakthrough" as a reference at 36 months, of 95 bits / 10 bits = **9.5 times greater** than Shannon's information, reflecting the lexical <information> with narrow meaning in the isolated words , thus allowing a quantification of the broad meaning by contextualizing the meaning in the sentence.
Therefore, the broad meaning is practically 1 order of magnitude greater than the narrow meaning, and therefore, with equal symbolic capacity and learning, it needs 10 times more time (as a learning

[24] Solé, R. (2009). *Redes complejas. Del genoma a Internet*. Tusquets Ed., Metatemas.

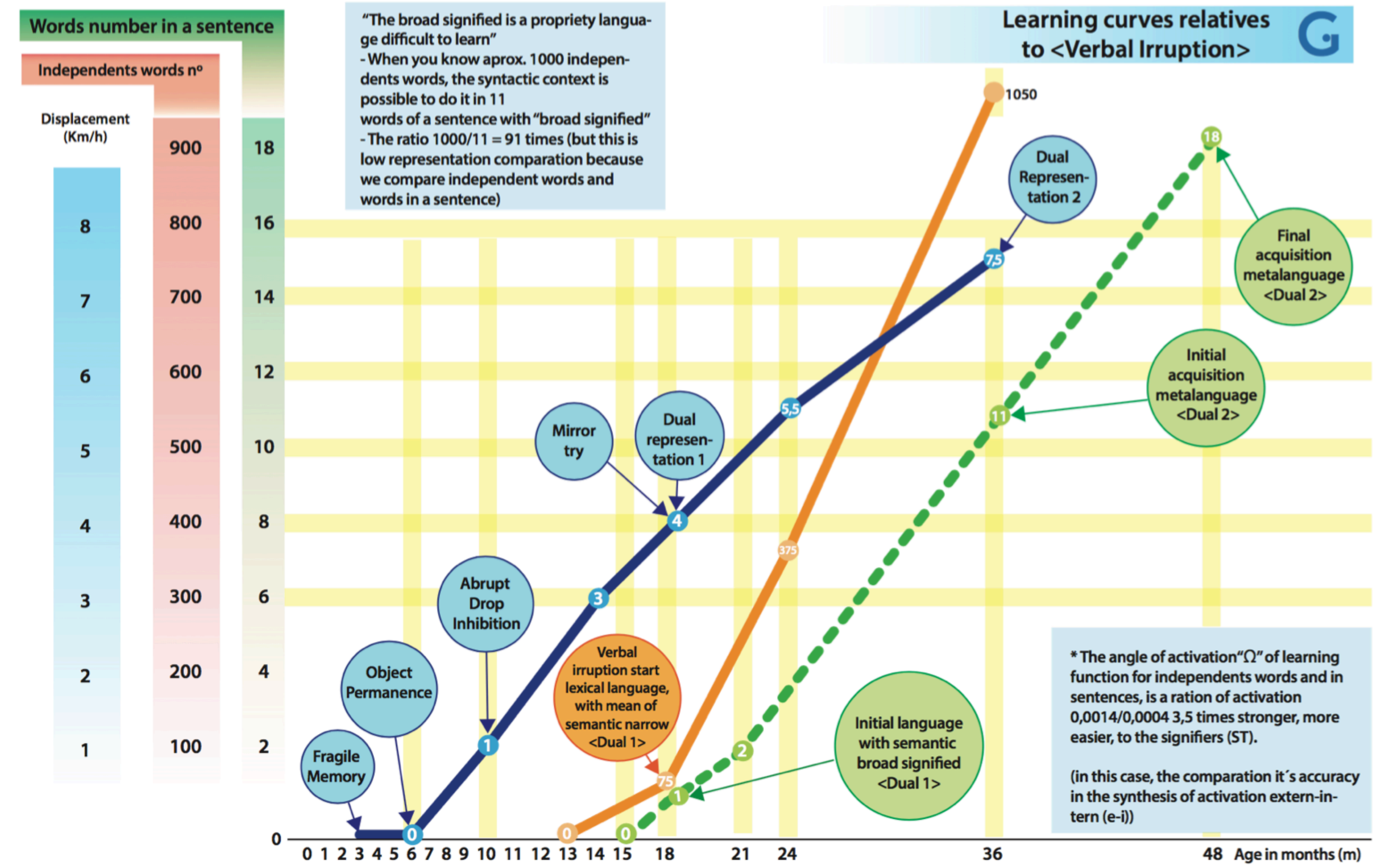
resource) for <cognitive accessibility>; or on the other hand, other resources to acquire meaning, which in the case of lexical vocabulary are lighter, in its meaning of narrow meaning.

**The signified**, for the linguistic case, and during the period of vocabulary learning and the use of a child's speech, is reasonably limited in its assessment of information in linguistic communication (the incorporation of audiovisual media, for example, allows a description parallel that can be analyzed, but that is not the case here).

**The additional dimension** that it supposes, represented in the complex plane (with complex numbering), means an increase in the capacity of the communication channel, to access meanings, in a 3-year-old child. This additional capacity, which is an order of magnitude greater than vocabulary (approximately x 10), and which unfolds in the developmental months from 13 to 48 months (orange and green curves in Annex 1), allow us to see its explanatory potential, practically completely unfolded.

Also the Semantic Receiver of the block diagram of Figure 2, represents this **new dimension** in the information and communication model of C. Shannon, and that W. Weaver already knew how to advance, that of signified.

**Annex 1**




www.ingit.es , in the Developing tab, Neuroscience and Psychology section, Article 4

- From these graphs, the graph of Figure 9 is deduced, specifically combining the green and orange curves under the parameter of the months of average age that the child needs for its development.


Note

Jordi Sos, a Telecommunications Engineer and lover of Quality as a criterion of excellence, is greatly appreciated for his contribution to the continuous bibliographic search, sometimes a repetitive request on my part, and always attended to. Thank you.