



# 5.1 TEORÍA MATEMÁTICA DEL SIGNIFICADO

Benigno Moreno Vidales<sup>1</sup>

Barcelona, 15/06/2021

## Índice

### - Introducción

- 1- Teoría matemática de la comunicación. W. Weaver (1949).
- 2- Teoría semiótica de la información. Paradigmas digital y analógico. P. Rocchi (2010).
- 3- Desarrollo de la teoría matemática del significado (TMS).

## **Introducción**

La piedra angular que supuso “La teoría matemática de la comunicación”, TMC (C. Shannon, 1948), permitió un avance muy importante en el campo de la transmisión de información.

Se abordó como un tema de ingeniería de las telecomunicaciones, y si bien N. Wiener también lo trabajaba desde la disciplina por el fundada de la cibernética, la diferencia fundamental era el carácter de la señal a transmitir, que en el caso de Wiener era “analógica” o continua.

En el caso de Shannon, las señal se trata en los términos que hoy denominamos “digitales”. Pero no es solo esta la diferencia, pues el término <información>, solamente incluye para el, a <la señal física> que se transporta, que transfiere para ser comunicada, y no al contenido semántico.

Intencionalmente C. Shannon no referencia en su propuesta, aquello que entendemos en lenguaje coloquial por “información”<sup>2</sup>, que evidentemente conlleva una señal física (auditiva, escrita, etcétera), así como su <significado semántico>, y el decide no contemplar esto en su propuesta, como decíamos.

Por tanto C. Shannon, a nivel lingüístico, se ocupa de la señal que lleva lo léxico, únicamente en su naturaleza <significante léxica>, dejando aparte el aspecto de contenido que ha de asociarse al <significado semántico>.

C. Shannon, trabajando en la Bell Telephone, y profesor del MIT, adopta el que su teoría vaya acompañada por un texto del veterano asesor de la Sloan Foundation, Warren Weaver.

En este texto, que figura como introductorio en la publicación de la TMC en 1949, sí se hace referencia extensa al <significado semántico>. Su título es: <<Contribuciones recientes a la teoría matemática de la comunicación>>.

Lo trataremos en el Apartado 1, de este estudio analítico, con una propuesta de modelo sobre la cuantificación del significado que Weaver apunta a lo largo de sus Contribuciones.

El aspecto de la teoría de C. Shannon se diferencia de la perspectiva de N. Wiener, diríamos en términos de hoy, como un modelo en el que se trabaja respectivamente con una señal “digital” discreta, y otro, con una señal “analógica” continua, como arriba apuntábamos.

La diferencia crucial entre ambos, es que el primero trabaja en su concepción fraccional o discreta, con código binario numérico para la transmisión de señal muy precisa, en el proceso de comunicación. El método de la “atomización” digital, permite controlar la precisión, exactitud y fiabilidad de la señal, y

---

<sup>1</sup> Referencias en [www.ingit.es](http://www.ingit.es). Registro de la Propiedad Intelectual <B-1377-21>

<sup>2</sup> “información”: en sentido normal o vulgar, que abarca a la “señal significante” y a su “significado semántico” conjuntamente.

también el poder pasar a una evaluación de la señal de <información><sup>3</sup>, generalización que conlleva el acto de comunicación, con el atributo que supone una libertad de elección de símbolos consecutivos. Un mensaje en el marco de tal evaluación, es el de la teoría de la probabilidad, en el cual se ha de elegir el símbolo a transmitir, que se configura con un cero (0), o bien un uno (1), su conjunto (0-1) es **un bit** de <información>.

Más allá de Shannon y Wiener, P. Rocchi (2010), recapitula unos 25 modelos de información hasta la actualidad (hasta el 2010), con el resumen y análisis más amplio conocido hasta fecha de hoy. Y acaba adoptando desde la lingüística un modelo semiótico, que “cobra” su propio peaje al abscribirse únicamente en términos de ese modelo digital, y que aunque realiza una generalización hacia los sistemas simbólicos, no es completa, aunque el lenguaje sea para nosotros (y por tanto su disciplina analítica, la semiótica), el instrumento simbólico por excelencia.

No obstante la importantísima síntesis y elección adoptada por Rocchi (Demming & Bell, 2012), permite tratar además, tanto el significante-léxico, como el significado-semántico, y por otro lado, esta generalizado al terreno de las telecomunicaciones, usando las terminologías de información “analógica”, así como el gran salto a la información con tecnología “digital” (de la que se hace conversión a “analógica” cuando se necesita; el “bypass” entre ambas, y en metodología matemática sistemática lo establecen métodos como el de Runge-Kutta, 1900), siendo clarificador informacionalmente, y un avance vital en la especialidad de las telecomunicaciones y la informática.

La visión de “significado”, se lleva hasta lo que luego se ha designado como “*significado estrecho*” (el que proporciona una palabra o término, sin desambiguar y “a secas”, una palabra aislada). Hinzen y Poeppel<sup>4</sup> (2011), apuntan al “*significado amplio*”, el de un término o palabra, dentro del contexto sintáctico de una frase (y anteriormente Frege, Putnam, Bearling y muchos más, en el sentido de identificar el significado-semántico como una característica autónoma y objetiva, y no sólo con atributos subjetivos, que es lo más habitual; “Meaning of the meaning” es buena referencia del propio Putnam como autor).

Por nuestra parte (B. Moreno, 2017), nos ha hecho falta para abordar ahora *el significado*, pasar por el concepto de Representaciones cognitivas (RC), dentro del Sistema Tridimensional de la Cognición (STC)<sup>5</sup>, dentro de la Psicofísica disciplinar (con Estímulos -ES-, las Sensaciones -SS-, y las Representaciones cognitivas -RC-).

Posteriormente hemos abordado la Capacidad simbólica con caracterización en la fisiología del sistema nervioso<sup>6</sup> (con correspondencia en el los componentes C1, C2 y C3, del ACP -análisis de componentes principales-, que ha sido realizado para establecer el STC), prolongando hasta el presente estudio de análisis matemático del significado.

El marco del significado-semántico, se aborda con formato matemático desde la perspectiva de W. Weaver, con procesos estocásticos de probabilidad, consideraciones a nivel entrópico de la información física, y sucesiva valoración semiótica, y en última etapa con P. Rocchi (2010), desde la señal, tanto analógica como digital, y con el atributo del significado-semántico sin cuantificar en su análisis.

---

<sup>3</sup> <información>: señal física, no se plantea su contenido. Lingüísticamente es el “significante léxico”.

<sup>4</sup> 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.

<sup>5</sup> En [www.ingit.es](http://www.ingit.es), en sección Proyectos, apartado de Neurociencia y Psicología, tesis para master en investigación en cognición y conducta: **3- Interacción Sensorial-Cognitiva en el Dominio Visual**

<sup>6</sup> En [www.ingit.es](http://www.ingit.es), en sección *En desarrollo*, apartado de *Neurociencia y Psicología*, dosieres: 1 Capacidad simbólica, 2 Infografías, Artículo 3 y Artículo 4.

Nuestra aportación de la cuantificación, primero del significado y su teoría, y luego con la aplicación sistemática en el <Tercer sector social y económico>, realizada en el caso particular del Colectivo de Personas Mayores (explícitamente en el artículo 5.2, siguiente al presente 5.1).

Se busca en esa aplicación, facilitarles la máxima <accesibilidad cognitiva>; se hace buscando pasar por una parrilla específica de comprobaciones, que modele la información con significado-semántico, más ampliamente que en la actualidad multimediativa, que siendo con cierta frecuencia “asignificativa”, es insuficiente para el objetivo aquí planteado para apoyos en ese Tercer Sector a personas con necesidad de apoyo cognitivo.

La parte que seguirá al presente estudio (5.1), que será como hemos dicho referenciada como 5.2, se investigará el/los cauces para la ampliación sistemática de <accesibilidad cognitiva> al significado amplio. El objetivo es que sirva a personas desfavorecidas cognitivamente en el sector de actividad referenciado arriba, y en la comunicación de todo tipo después, incluidos los discapacitados intelectuales, y en general en los procesos de aprendizaje: en la educación, formación y entrenamiento, generalizando hasta la divulgación general y a la comunicación en los medios de comunicación de masas en nuestra sociedad del conocimiento.

## 1- Teoría matemática de la comunicación

En <Contribuciones recientes a la teoría matemática de la comunicación>, W. Weaver (1949), establece una definición muy psicológica, “(...) *comunicación (...), como el conjunto de procedimientos por los cuales una mente puede afectar a otra*”, tanto oralmente como por escrito, con música, con pintura, teatro, ballet y en general todas las manifestaciones humanas.

Los niveles de comunicación por él propuestos son:

- Nivel A: como problema técnico de la transmisión de símbolos en la comunicación.
- Nivel B: como problema semántico de ser recibidos los símbolos con el significado deseado.
- Nivel C: como problema de efectividad del significado para afectar con una conducta deseada.

“A” es tratado por la teoría de la comunicación de C. Shannon.

Nosotros vamos a centrarnos en “B” en este estudio 5.1, y estudiamos la identidad o aproximación del significado captado por el receptor, comparando con el significado previsto por el emisor o fuente de información.

El problema del nivel “C” de la efectividad, esta relacionado con el problema semántico de A y B, por un lado, y por otro se tratará en el artículo aparte que referenciamos como 5.2, como se ha avanzado.

W. Weaver determina que los niveles B y C, son dependientes del A, por que este incluye parcialmente a los dos niveles siguientes. Aún así Shannon se ratifica en que “*los aspectos semánticos de la comunicación son irrelevantes desde el punto de vista de la ingeniería*”.

Y a su vez, Weaver matiza “*pero esto no quiere decir que los aspectos de ingeniería sean irrelevantes desde el punto de vista semántico*”.

Según las notas 2 y 3 de la página 1 de nuestro estudio, el término <información> en la teoría de la comunicación se refiere, no tanto a “lo que se dice”, como a “lo que se podría decir”, y añade, “**la <información> es la medida de la libre elección de un mensaje**”.

Entre dos mensajes posibles a enviar, subyace la elección de uno. El conjunto de los dos, es la unidad de información, y se puede pensar que uno de ellos lo es, pero esto es un error flagrante. Es el conjunto de los dos, la unidad de medición que rige/constituye a la <información>, que es el Bit (*Bynary Digit*)<sup>7</sup>.

Por tanto en su base más simple, irreductible, es una elección binaria, un <<bynary digit>> (0-1), y que aritméticamente los representa la expresión de potenciación:  $2^1$  para una simple elección 0-1, o un bit;  $2^2$  para una elección entre 4 opciones posibles, o dos bits;  $2^3$  para una elección entre 8 posibles, o tres bits; y así sucesivamente, por lo que la medida de un conjunto de elecciones de,

- 2 casos, es el logaritmo en base 2,  $\log_2 2 = 1$  bit
- 4 casos, es el logaritmo en base 2,  $\log_2 4 = 2$  bit
- 8 casos, es el logaritmo en base 2,  $\log_2 8 = 3$  bit
- ... etcétera

Siendo así la medida logarítmica *totalmente natural*, en cuanto a razón de estar aritméticamente tipificada, con una lógica electiva.

Tomemos las 27 letras del alfabeto inglés como la posible elección entre 27 mensajes, supongamos siendo equiprobables cada uno de las 27 letras / símbolos (que no lo son exactamente, pero haciéndolo por simplicidad de cálculo y una aproximación relativa), la información es de un valor,

- $\log_2 27 = 4,76$  bits

Si a cada letra/símbolo, le atribuimos su probabilidad de uso en inglés  $p_A$ ,  $p_B$ ,  $p_C$ , ... etc, cuya suma es 1, podemos establecer que para cada letra, tomando su probabilidad en "bits", es la suma sucesiva,

- $p_A \log_2 p_A + p_B \log_2 p_B + p_C \log_2 p_C \dots$  etc
- 

(siendo  $p_A$  la porción probable de A, y  $\log_2 p_A$  los bits de A, y así sucesivamente)

Cuya información en este caso, con la precisión de ocurrencia probable de cada letra (símbolo), que es conocida por su uso estadístico, por ejemplo en la lengua inglesa, la información que aporta realmente es una media de **4 bits**. Esta media es menor que los 4,76 bits con equiprobabilidad, visto en la fórmula y cálculo primera.

Este cambio se debe a la <redundancia>, que es imprescindible para que un mensaje sea inteligible entre ruidos que aparecen durante la transmisión, acotándose por la repetición que desambigua.

Esta cifra es aún menor si hacemos el cálculo ensayando secuencias<sup>8</sup> de una frase, dándonos una cota todavía más baja, de **2 bit**, del contenido informativo real del alfabeto inglés.

Si la libertad de elección máxima es 4,76, y la real es 2, la relación de las dos medidas es la libertad relativa:  $2/4,76 = 42\%$ , por lo que *la redundancia*, sería  $100 - 42 = 58\%$ .

Como comparación en otros lenguajes, cada símbolo tiene un valor informativamente hablando:

- El código Morse (punto- raya): **1 bit** (que es el  $\log_2 2$ )
- El código de numeración decimal (0, 1, 2, 3,...etc): **3,32 bit** (que es el  $\log_2 10$ )

Por tanto generalizando la fórmula de los símbolos del alfabeto a los "n" símbolos de cualquier posible mensaje entre la colección de esos mensajes, su valor generalizado es:

---

<sup>7</sup> W. Tukey usó el término Bit por primera vez. No confundir bits con bytes (combinación de normalmente 8 bits)

<sup>8</sup> J. Singh (1966). *Teoría de la información del lenguaje y de la cibernética*. Alianza Editorial 1972.

-  $\sum_1^n p_n \log_2 p_n$  (en bit's)

La denomina Shannon <<H>>, y como  $p_n$  es una fracción de 1, el logaritmo será un cantidad negativa, para su expresión mejorada en positivo, con lo que la información H, se expresa como:

-  $H = - \sum_1^n p_n \log_2 p_n$  [1]

Conceptos como el ruido, redundancia, uso de bits de paridad para asegurar la fidelidad del mensaje, y otras técnicas y sistemas que lo acompañan en su transmisión, merman a su vez la capacidad del canal para transmitir símbolos (mensajes) por unidad de tiempo, pero lo hacen en forma más segura. En este sentido nos interesa la visión del sistema en un diagrama de bloques, no entrando en detalles (que son vitales para la buena transmisión, pero no necesarios aquí). La comunicación en diagrama de bloques, es la siguiente,

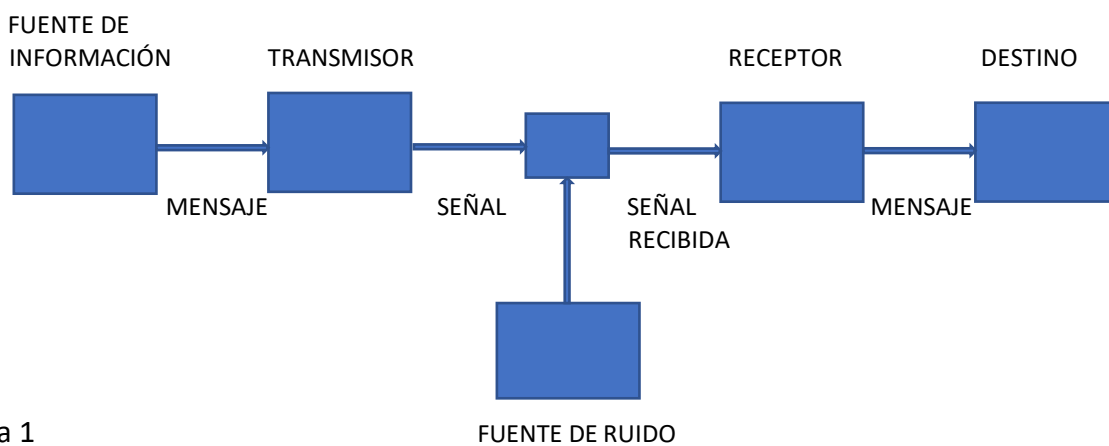


Figura 1

Siguiendo con la contribución de Weaver, su sugerencia (1ª) respecto al diagrama de bloques anterior, y referido a tratar los niveles de comunicación B y C a continuación, esta revisión consiste en añadir en el diagrama otra caja que se denominaría “Receptor Semántico”, colocada entre el Receptor de ingeniería que cambia las señales a mensajes, y del Destino. Este receptor semántico realiza una segunda decodificación del mensaje.

Igualmente se introduce otro bloque entre la Fuente de Información y el Transmisor, referenciándolo como Ruido Semántico, dejando para la fuente de ruido general la denominación de Ruido de Ingeniería. Lo podemos ver en la Figura 2.

Una sugerencia (2ª) de Weaver, es que conviene estudiar la estructura estadística del lenguaje, siendo partidario de utilizar la teoría de las cadenas de Markov, al adaptarse a uno de los mas importantes y difíciles aspectos del significado, la influencia del contexto. En nuestro trabajo aquí, se ha adoptado este punto de vista al considerar la palabra como el primer componente del lenguaje, y a la frase (como cadena de palabras) como el contexto para dotar de más significado explícito a cada palabra. Cómo átomo del lenguaje, la palabra es la entidad que las reglas sintácticas no pueden separar o reorganizar.

La siguiente sugerencia (3ª), expresa que “Se tiene el vago sentimiento de que la <información> y el significado pudieran tratarse como un par de variables canónicamente conjugadas (...)”, aspecto que aquí, en esta teoría, es ampliamente analizado y mostrado en el Apartado 3.

La siguiente sugerencia que se analizará (4ª), se refiere a *la entropía* introducida por Shannon en su teoría como paralelismo a *la información*, considerada como que “(...) *es con seguridad el hecho más importante*” en el nivel “A”. En el nivel de información B y C, se volverá a poner su relevancia de manifiesto. Se verá en el artículo 5.2, continuación del presente 5.1.

Por último (5ª) manifiesta que “(...) *la entropía no solo habla del lenguaje de la aritmética, sino también del lenguaje del lenguaje*”. Efectivamente en esta superposición, el metalenguaje, abre la puerta a como surge esta característica del desarrollo infantil en la que después de la aparición de la diferenciación simbólica (Representación Dual 1, según J. Delohache), aparece una segunda perspectiva metarepresentacional, radical para todos los procesos de información en sentido amplio. En la curva verde del Anexo 1, observamos inicio y final a los 36 y 48 meses de la Representación Dual 1 y 2, motivo de la habilidad de las metarrepresentaciones y del metalenguaje. También se tratará exhaustivamente en el artículo 5.2 comentado.

### Detalles de las cinco observaciones anteriores

Los 5 puntos aducidos por Weaver los suscribimos mayoritariamente. Todos ellos representan parte de las propuestas del presente estudio, y que vamos a examinar con esas matizaciones y complementos,

(1ª) Diagrama de bloques ampliado con receptor y ruido semánticos, el diagrama de bloques de la comunicación, quedaría así:

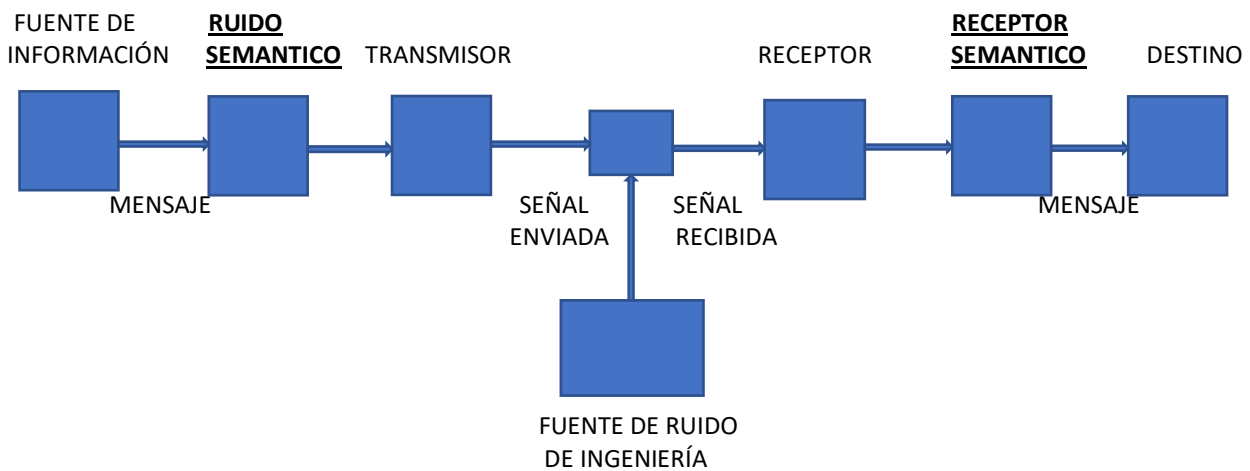


Figura 2

Para concretar el marco de referencia del modelo, hemos elegido el proceso emergente en el desarrollo infantil, de la <irrupción verbal> en bebés-niños a partir de los 18 meses<sup>9</sup>, con sus curvas de aprendizaje que veremos en el apartado 3.3, en las que aparecen tanto de vocabulario léxico (significante-léxico), como de vocabulario con significado amplio (significado-semántico). En su análisis se pone de manifiesto que si la primera tiene un coste de un nivel 1 de aprendizaje, la segunda tiene un coste en recursos de 3,5 para el mismo sujeto, para el aprendizaje del significado de su universo de vocabulario.

Al introducir las cajas en el diagrama de bloques, de Ruido Semántico y de Receptor Semántico, se están significando estos costes diferenciales para el proceso de comunicación que supone la

<sup>9</sup> En [www.ingit.es](http://www.ingit.es), sección *En desarrollo*, apartado *Neurociencia y Psicología*, Artículo 4, ver curvas crecimiento.

“irrupción verbal”, y que se produce entre el medio estimulante cultural y la mente del niño, formado principalmente por su madre y familia, y luego por toda la red social próxima (en la guardería, el parque de juegos, etcétera).

Esto sucede en una realimentación continua de estimulación en la ventana de edad desde su nacimiento hasta los 48 meses, y que en caso de que no se produjera<sup>10</sup> esa interacción, puede provocar lesiones permanentes como el caso de Victor de Aveyron (criado como “niño salvaje” en el bosque sin contacto humano) o el de niños en los orfanatos de Polonia durante la segunda guerra mundial, sin trato personal alguno.

(2ª) La frase. El lenguaje con palabras (*unidad mínima* del lenguaje, Bickerton<sup>11</sup>, 2000) y con las frases (como combinaciones de esas unidades mínimas), se establecen como referentes para la evaluación de la <irrupción verbal> en bebés-niños.

La palabra significante-léxica, inserta en una frase, requiere por parte del bebe-niño un conocimiento del significado-semántico de cada término, mucho más profundo en su sentido, y en función de la longitud de la frase medida en número de palabras, y que crean un contexto significativo.

En esta propuesta de Weaver, se alude al análisis de frases que contextualizan a la palabra independiente, en una cadena de Markov. En nuestro caso, entramos a valorar directamente la frase con ese potencial de contexto de las palabras, y nos centramos en su aspecto puramente semiótico y lingüístico, y lingüísticamente una oración, mas que una cadena, es un árbol (estructura sintagmática).

Además está el trasfondo sintáctico puro, configurado según la Gramática Generativa Universal de N. Chomsky, con aproximadamente 24 reglas-funcionales comunes de composición, existentes en todas las lenguas, donde se estructura esa articulación de una frase con sentido o significado. Por otro lado, partimos del principio, de que en el lenguaje como agente simbólico cultural por excelencia, en el que “lo aprendido se hace instinto” según el <efecto Baldwin>, llevando tras de sí todo el proceso evolutivo biológico<sup>12</sup>, así como antropológico que estructura el lenguaje, y que se manifiesta en el desarrollo de cada bebe-niño, inmerso en la cultura, con influencia probablemente decisiva en los aspectos genéticos y epigenéticos<sup>13</sup>.

El conjunto de factores establece un significado estrecho en cada *palabra*, y uno contextual y amplio (Hinzen y Poeppel, 2011), respecto la palabra independiente, que está fuera de una *frase*, *clausula*, *párrafo*, *texto*, etcétera.

(3ª) La unión arbitraria del significante-léxico y el significado-semántico, es un indicador de independencia de ambas simbolizaciones (F. Saussure, 1928).

Las dimensiones son el constructo matemático cartesiano, que permite la conjugación de las variables, en este caso la de <información> (en términos de Shannon) y el significado en un espacio bidimensional, algebraico y geométrico, como variables independientes en un plano bidimensional.

---

<sup>10</sup> En [www.ingit.es](http://www.ingit.es), sección *En desarrollo*, apartado *Neurociencia y Psicología*, Artículo 3, pags. 9, 25, 26 y 28.

<sup>11</sup> Bickerton & Calvin (2000). *Lingua ex Machina*. La conciliación de las teorías de Darwin y Chomsky sobre el cerebro humano. Editorial Gedisa. Pag. 37, 54 y 63

<sup>12</sup> Sampredo, J. (2002). *Deconstruyendo a Darwin*. Ed. Crítica s.l. 2002. – Ver nota 10, Bickerton, pág. 241

<sup>13</sup> Ver en [www.ingit.es](http://www.ingit.es), sección *En desarrollo*, apartado *Neurociencia y Psicología*, Artículo 3, pags. 10 y 26.



P. Rocchi (2010)<sup>14</sup>, expone su planteamiento del triángulo semiótico (Figura 3), como el modelo óptimo, utilizando dentro del modelo lingüístico, el significante (léxico) y el significado (semántico), por un lado, y por otro el polo referente observador, que puede ser una persona o un dispositivo instrumental detector-actuador que interacciona con los polos significante y de significado. Este último permite, siempre que lo evidencie el indicador de objeto-suceso (OS), tener una magnitud que representa en algún aspecto al OS, como un indicio, índice, indicador, nota, o como un icono, o simplemente simbolizando arbitrariamente al OS. Lo veremos en mayor detalle en el apartado 2 de este estudio.

(4ª) Pero esta “Teoría matemática del significado”, no llegaría hasta el final de las propuestas de Weaver, si no abordara la entropía como hilo conductor en el mundo de la física con su segundo principio de Termodinámica.

Desde las polaridades de Orden-Caos, se han desarrollado teorías de sistemas<sup>15</sup>, en algunos casos jerárquicos, muy estructurados y ordenados, así como teorías del caos<sup>16</sup>, estas últimas como polo de referencia para la entropía, del mismo modo que los sistemas en el otro extremo la tienen en la <información> (en sentido estrecho).

Justo en el centro<sup>17</sup> de estas dos polaridades, surge la Teoría de la Complejidad Adaptativa, basada en la autoorganización adaptativa emergente<sup>18</sup>, cuyo exponente quizá más sistemático es la Teoría de Redes Complejas<sup>19</sup>.

Un centro de estudios, como es el Instituto de Santa Fe, ha tenido profesionales de los campos de todas estas teorías: de *sistemas*, del *caos*, de la *complejidad*, que han aportado conceptos y planteamientos esenciales, y que pivotan en nuestra opinión, sobre la física de la segunda ley de la termodinámica, la de la entropía.

En el Instituto de Santa Fé, desde Laughton a Kauffman<sup>20</sup> y a Gell-Man; el físico A. Solé, de la U. De Barcelona que estuvo también en el instituto en una 2ª época, siendo por otro lado quizá, el más notable alumno del físico y museólogo Jorge Wagensberg. Ambos en la facultad de físicas de la UB, tuvieron una interesante relación con Margaleff, el relevante ecobiólogo de la UB que aplicó la teoría de la información para cálculos de los nichos biológicos, al que siguió J. Terradas en la misma línea y que actualmente continúa F. Lloret, especialistas en biodiversidad.

A. Solé, con sus trabajos en Teoría de Redes, con aplicaciones en lingüística, y en general de los sistemas complejos, atrae especialmente la atención de este estudio multidisciplinar aquí presentado.

Probablemente las redes complejas, y en particular los trabajos de Erdos y Renyi, configuran el espacio en que converge la teoría de la complejidad adaptativa autoorganizada, con una fortísima componente con la estocástica matemática muy amplia (como la de las cadenas de Markov). Por supuesto, su relación con la lingüística, es porque prolifera en estructuras superficiales y profundas (estructuras “s” y “p”), con árboles de sintagmas, con grandes niveles de complejidad. La sintaxis es compleja, pero su complejidad tiene una razón, y es que el pensamiento es más complejo que el lenguaje, y solo podemos expresarlo a través de un métodos limitados, como el lenguaje de palabra en palabra (en oraciones, que son microcosmos de palabras).

---

<sup>14</sup> Rocchi, Paolo (2010). *Logic of Analog and Digital Machines*. Pags. 7, 32 y 38

<sup>15</sup> Betalanffy, L. v. (1969) .Teoría de sistemas

<sup>16</sup> Gleick, J. (1987; 2011). Dos textos: *Caos e Información*

<sup>17</sup> Gell-Man, M. (1994). *El quark y el jaguar*. Tusquets Ed.. Pags. 77, 123

<sup>18</sup> Johnson, S.(2001). *Sistemas emergentes*. Turner Fondo de Cultura Económica

<sup>19</sup> Solé, R. (2009). *Redes complejas. Del genoma a Internet*. Tusquets Ed., Metatemáticas.

<sup>20</sup> Kauffman, S. (2000). *Investigaciones*. Tusquets Ed., Metatemáticas.

La gramática ofrece argumentos para la refutación de la doctrina empirista, de que no existe nada en el intelecto que no haya pasado antes por los sentidos.

En esta disciplina, la *Emergencia autoorganizada adaptativa* en el Borde del caos entre información y entropía, es donde encontramos fractales (estructuras arborescentes para el lenguaje), la organización de muchos procesos biológicos, y la tensión de “convivencia” entre la <entropía> y la <información>.

El excurso que hacemos aquí es muy amplio por la importancia, a nuestro juicio, del objetivo final que acabando su texto, W. Weeber expresa junto a la relevancia de la entropía para el tema de significado, y es el del *lenguaje del lenguaje*.

(5ª) El metalenguaje es vital en el campo del desarrollo infantil (Perner, 1991)<sup>21</sup>, por lo tanto además del trato de la <información>, la entropía y el significado como Nivel B de la comunicación en la clasificación de Weeber, se manifiesta el criterio esencial del metalenguaje.

El metalenguaje permite la reflexión en el pensamiento (entre otros, lo trata en su Fenomenología, Husserl), permitiendo la recurrencia tanto en el pensamiento reflexivo como en el lenguaje hablando sobre si mismo.

Es el origen semiótico tratado por U. Eco (1975) como “factible de mentir” (desentrelazamiento, entropía del significante-significado), que se inicia en el desarrollo del niño a los 36 meses y termina a los 48, como el de la representación dual 2 de J. Delohache. Mentir supone desentrelazamiento, entropía del significante-significado.

También crea paradojas, que se resuelven también desde el mismo metalenguaje. Desde el teorema de Tarsky se hace desanudando la “paradoja del mentiroso” de Epiménides, y mucho más allá al establecerse la verdad matemática de la <incompletitud> por el teorema de Gödel. En él se demuestra que cualquier sistema simbólico es incompleto si es consistente.

El campo común del teorema de Tarsky y el de Gödel, es en nuestra opinión el metalenguaje, cuya base está en el lenguaje modelado por la polaridad de los *significantes – significados* (en términos de Weeber [<información> - significado], tema central de este estudio).

Pero sin “significado amplio semántico” y permanentemente abierto a expandirse, sobre cualquier objeto-suceso (OS), que apuntemos, que se dese “narrar” en términos, no hay información conjugada con el significado.

El metalenguaje por su autorreflejo, permite disponer finalmente de un marco teórico y matemático de la consciencia reflexiva animal y humana, que puede encuadrar enfoques para la inteligencia artificial (que hoy, en parte, está cautiva de los efectos referenciados por la “habitación china” de J. Searle, del “hard problem” de Chalmers, la brecha separadora del “cuerpo-mente” de Levine, o/y el “grounding problem” de Harnad).

Esta parte se desarrollará como se ha comentado, en un Artículo 5.2, continuidad del presente.

---

<sup>21</sup> Perner, Josep (1991). *Understanding the representational mind*. MIT

2- Teoría semiótica de la información, según la <<Lógica de las máquinas analógicas y digitales>> de P. Rocchi, 2010.

P. J. Denning, hace el prólogo de este texto con la siguiente pregunta de P. Rocchi “¿Qué es la información?”. Más tarde en un artículo en “American Scientist”, firmado con Bell, abordaran la singularidad de la respuesta de P. Rocchi (2010) a aquella pregunta, y fue: es la información según C. Shannon, y es el significado<sup>22</sup>, dentro del triángulo semiótico “significante – significado - referente”; y que esta, está asociada a la información significativa o <información> de Shannon; y se manifiesta lingüística y semióticamente en la interpretación del Referente (receptor o emisor <R>, que trabaja con un *indicador o indicio*, o un *icono*, o/y un *símbolo* del objeto-suceso OS *representado*, motivo de la comunicación).

Está relación se expresa desde el símbolo como una representación **artificial**

La apreciación semiótica y lingüística nos permite ver que “La palabra es la **quintaesencia del símbolo**” (S. Pinker, 1994)<sup>23</sup>, nos enmarca la naturaleza del símbolo y la palabra, como correspondientes.

El “símbolo” es el término usado por Turing para plantear el componente unidad de su máquina universal de procesamiento (primera propuesta de ordenador), también empleado por Shannon en la TMC, en inteligencia artificial por Pitts & McCulloch, en una larga tradición de autores en disciplinas cuando tratan la información (Gadner, Goodman, Deacon, Eco, etcétera).

Del símbolo, y de la palabra en particular, los niños han de asumir tácitamente que *no son notas del comportamientos subjetivo* de una determinada persona, sino que las mismas se comparten de forma bidireccional, y que al que habla, le sirve para transformar significados en sonidos, y al que escucha, los sonidos en significados, usando ambos el mismo código.

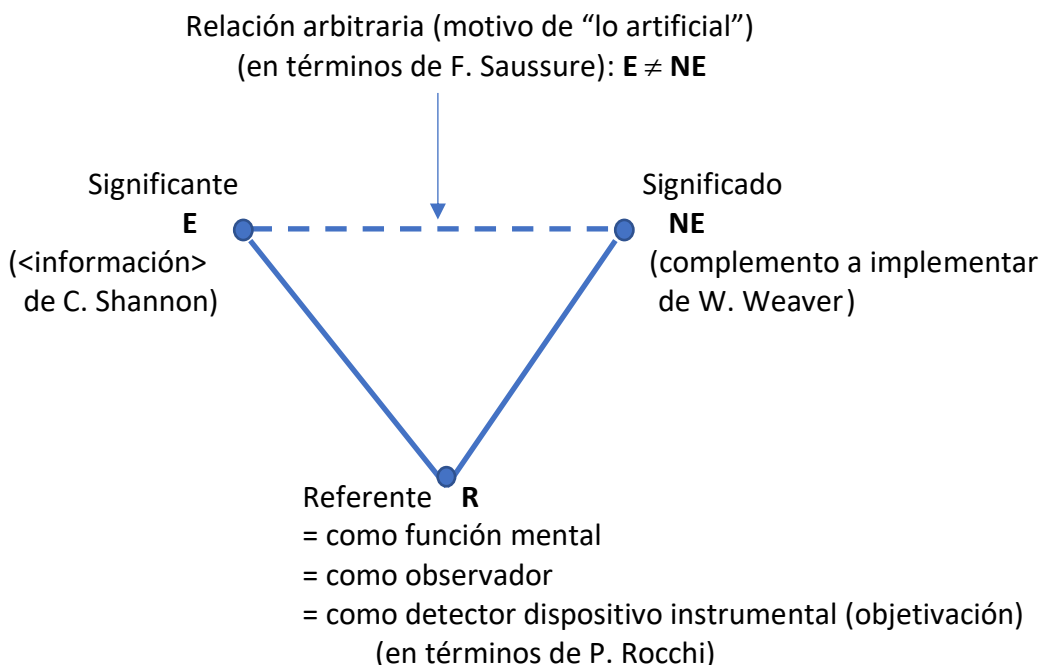


Figura 3

<sup>22</sup> Denning & Bell (2012). *The information paradox*. American Scientist Nov.-Dec. 2012

<sup>23</sup> Pinker, S. (1994). *El instinto del lenguaje*. Alianza Editorial

El símbolo convertido en un círculo desde el triángulo anterior, por identificar E = NE surge como una representación-objeto de los OS *natural*

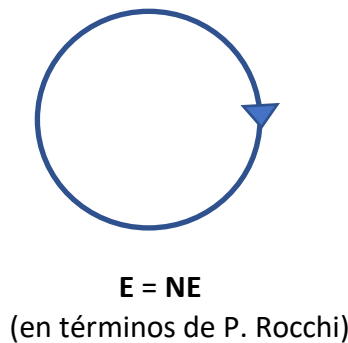


Figura 4

P. Rocchi, radicalmente riguroso en el examen de la semántica, expone realizar “(...) un intento independiente de diseccionar las tecnologías de la información sobre la base de propiedades físicas y conceptos semánticos también” (pág.40). El subrayado es nuestro.

Y sigue, “*Los factores subjetivos interfieren con la semántica*”, no pareciendo obvio por este factor, dejar a un lado la semántica. A continuación cita tres campos del conocimiento que se ha abordado fructíferamente a pesar de la subjetividad implícita que tienen (como la semántica):

- 1) La probabilidad p(A), en los métodos bayesianos de la escuela subjetivista con p(A), que expresa el grado de creencia de un individuo sobre la ocurrencia de A.
- 2) Los sistemas de referencia en el movimiento de la física cinemática, en los que la velocidad de un cuerpo depende del observador, en el sentido de que dos observadores pueden asignar valores de velocidad muy diferentes (p.e. movimiento de un tren en la vía paralela respecto a otro tren, ¿quién se mueve, él o nosotros?).
- 3) La medicina, como disciplina personalizada; por ejemplo, de dos pacientes tratados con el mismo protocolo, a veces difieren de manera drástica, cuando uno muere, y otro sobrevive (también p.e. por la subjetiva tolerancia a la farmacopea y a todos los remedios médicos; y por esta única razón, no llevó a los expertos a abandonar el estudio de la medicina).

Continuando: “*Los desafíos científicos antes mencionados no constituyen curiosidades intelectuales o diminutas trayectorias pertenecientes al mosaico científico*”.

“*Los científicos no pueden destruir el subjetivismo de los fenómenos naturales, sin embargo no se rinden y trabajan duro para mejorar el conocimiento de los fenómenos y delimitar los efectos de la arbitrariedad”.* No podríamos estar más de acuerdo.

“*No veo razones para eludir esfuerzos paralelos en la ciencia de la información, y este libro representa un intento de mostrar como la semántica puede discutirse sobre la base de problemas objetivos (...)”.*

Lo digital puede reconstruir lo analógico, pero lo inverso no es cierto, por ser más borroso, menos “nítido” (principio de nitidez de P. Rocchi). Pero hay que decir que lo analógico es lo que manejamos, como organismos y humanos, más habitualmente.

Por ejemplo, con **el lenguaje**, el proceso simbólico más relevante con el que nos manejamos, tiene sin embargo una componente fraccional (si no digital), notable: letras, palabras, frases, párrafos, textos, etcétera. Pero esto no evita que para escuchar y emitir las palabras, tengamos que usar el “medio aire” con una señal acústica, intrínsecamente analógica.

Rocchi establece una sucesión aún más progresiva en niveles del 1 al 5 para la estructura informacional:

- 1) Bit's
- 2) Palabras binarias
- 3) Palabras comunes
- 4) Textos (e imágenes y sonidos), e
- 5) Información multimedia e hipermedia

La migración de la información más fraccional a la más holística, sigue la trayectoria también de lo más digital a lo más analógico, y de lo más significativo a lo que tiene más significado (por ejemplo desde el nivel 3 al 4, para pasar de las palabras a los textos con oraciones, en la escala referenciada).

Esta progresión, Rocchi la denomina <<ensamblaje>>, y tiene una gran similitud conceptual con lo que hemos aquí denominado <entrelazamiento>, para caracterizar la relación de las dos dimensiones diferentes, pero íntimamente relacionadas, de significativo y significado, que se vera en el Apartado 3.

No es un término de construcción conceptual desmesurado. Dice *“una página web que ofrece textos, piezas musicales, mapas, diagramas, películas y herramientas interactivas constituye un buen ejemplo de comunicación hipermedia”*.

### **2.1 <El significado>, como un término progresivo, incremental e ilimitadamente factible de ser creciente de sentido**

La tecnología se representa por las máquinas que la forman.

Las máquinas manipulan la <información> y el significado en forma siempre ampliable.

Permiten entender un paso más en órganos como es el cerebro y su función mental, y la codificación; así mismo, también posiblemente en máquinas como los ordenadores y sus procesos, y su similitud con el proceso de pensamiento, con el aserto de Miller sobre la recodificación: <<me parece que es la savia vital de los procesos mentales>> (Gleick, 2011, p.364). Veamos varios dispositivos operativos en la informática,

#### **21.1 Convertidores**

211.1 Una tecla: hace la conversión analógica mecánico-eléctrica.

211.2 Un led: convierte los significantes eléctricos en luz.

211.3 Una impresora: convierte electricidad -> letra impresa

211.4 Una pantalla: bit's eléctricos -> pixels

211.5 Un altavoz: ondas eléctricas -> sonidos

211.6 Un CCD (cámara “charge coupled device): señal luminosa -> eléctrica

211.7 Un imán

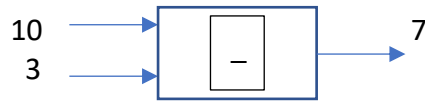
211.8 Un motor eléctrico

etcétera

En todos ellos, se produce la conversión de los significantes de unos a otros ( $E_{input} \neq E_{output}$ ), manteniendo *inalterados los significados* ( $NE_{input} = NE_{output}$ ); ver en 2.2.

21.2 Procesadores de datos (crean nuevos conjuntos de <información> términos, frases, contextos en textos, idiomas, etcétera).

### 212.1 Numéricos. Realizando cálculos numéricos



212.2 Verbales: Un ejemplo en Internet con la función “Búsqueda”, para encontrar “un teléfono móvil”, nos vamos a -> Webs con distintas ofertas de telefonía.

### 212.3 Visuales:

- secuencias de imágenes sucesivas, como los vídeos.

### 212.4 Simulaciones:

- de aviones en vuelo, en química, física, etc.
- extractos económicos
- noticias
- manejo de sonidos, textos, letras, imágenes

### 212.5 En lo real:

- análisis preciso del significado en la “experiencia práctica”, con números que lo representan (con una etiqueta)
- dólares
- euros
- litros
- productos ... etc

### 212.6 Creaciones:

- la máquina no inventa, pero tiene creatividad, proporcionando ideas originales, por ejemplo atribuyendo o deduciendo significado como de 212.1 a 212.5

### 212.7 Conversiones en nuevos datos y en la realidad:

- “¿cómo puede un sistema informático que es inconsciente y desconoce el significado de sus datos, transmitir información novedosa?”

⇒ Por instrucciones aplicadas a las bases de datos, desde la memorias de esos datos y estableciendo relaciones y correlaciones (por ejemplo entre el espacio recorrido, y el tiempo empleado en recorrerlo, nos da el ratio la velocidad, *información novedosa* no obvia desde los datos espacio-tiempo de partida).

### 212.8 Programadores de software

- pueden: artesanalmente manejar instrucciones y bases de datos, no de forma fija, pero sí creativa.
- no pueden: implementar una aplicación de “software eficaz” de una forma simple, con una metrología rígida.

## 2.2 En un ordenador se manejan E (significantes) y NE (significados), usando unidades distintas de operación en diferentes configuraciones

- Funcionando juntas E y NE (como vimos ya como introducción, en el apartado 21.1)

(1) Variando solo los significantes:  $(E_{input} \neq E_{output})$   
y manteniendo inalterados los significados  $(NE_{input} = NE_{output})$

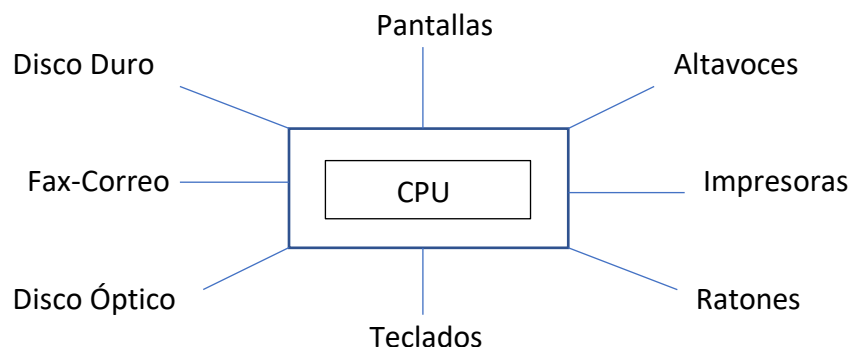
Siendo (1), por tanto un conversión *no informativa* (solo varían los significantes)

(2) Varían solo los significados:  $(NE_{input} \neq NE_{output})$   
y manteniendo inalterados los significantes  $(E_{input} = E_{output})$

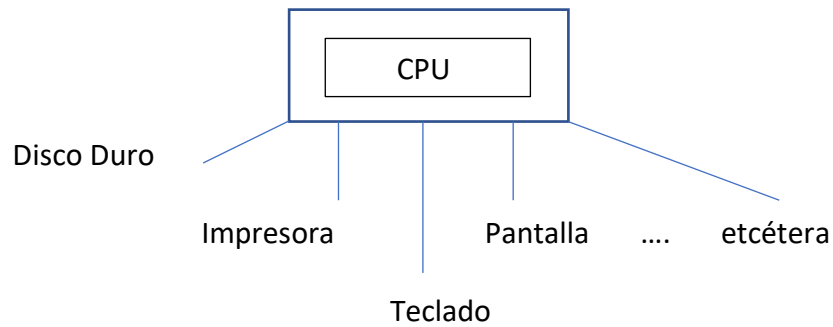
*Siendo (2) novedoso, **informativo**.*

Por la estructuras o arquitecturas de los dispositivos tecnológicos unidos entre ellos, definen las posibilidades de la <convergencia digital> en distintos formatos, así como con una variedad de **capacidad informacional** relativa:

22.1 En estrella (CPU, Control Process Unit): modelo radial que lo facilitan claramente las entradas USB (Universal Serial Bus) y la red de redes, Internet,

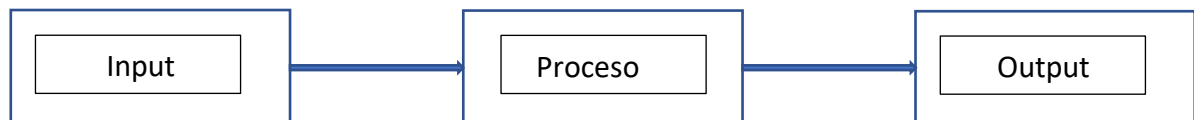


22.2 En árbol: modelo jerárquico comandado; como los teléfonos móviles, i-pods de alta fidelidad, controles de juegos, navegación de satélites,



### 22.3 En cadena (modelo IPO; Input, Process, Output):

Se ve frecuentemente en producción agrícola, hídrica, minería..., pero también en procesamiento de números, imágenes, sonidos ...



Es equivocación frecuente, en la disciplina Informática, el preparar los manuales técnicos como “libros de cocina”, cuando deben comunicar los <principios teóricos informáticos>, establecidos hoy en día preferentemente en el paradigma digital, tanto en sus bases, como en sus objetivos, no con un recetario, como se hace demasiado frecuentemente.

La CPU y su estructuración de periféricos, es la clave, pues su procesamiento, su capacidad combinatoria, articula en forma distinta la información procesada, como lo hace la fórmula de la velocidad al dividir el espacio por el tiempo (apartado 212.7)

22.4 El escalado niveles 1 al 5 (página 12): desde los significantes más simples, como lo son los Bit’s o las palabras binarias (en nivel 1), hasta indicadores con indicios perceptibles sensorialmente, como lo hacen los multimedia/hipermedia (en nivel 5); lo cual no es simple, y mucho menos, evidente.

En este nivel 5, es significativo *lo audiovisual*: desde el indicador-indicio mostrado en forma correlacionada con un texto, hasta las etiquetas sistemáticas<sup>24</sup>, o buenas correlaciones con textos narrativos, etcétera.

La “conversión” pura de código (según la fórmula **(1)**) no aporta significado-semántico; cuando intervienen “instrucciones” sobre una base de datos (según la fórmula **(2)**), podemos disponer de

---

<sup>24</sup> - Deng, Jai, y otros (2009). *ImageNet: A Large-Scale Hierarchical Image Database*. Dept. of Computer Science, Princeton University, USA. En la Universidad de Princeton: más de 45.000 imágenes etiquetadas temáticamente.

- Von Ahm, L. & Babbish, L. (2004). *Labeling images with computer game*. ACM Press, pp.319-326



un output del conjunto que aporte nuevo significado, formando significado-semántico: por combinatoria, por agregación, por sintaxis, etc.

En conclusión, y también a modo de resumen aquí sobre “Logic of Analog and Digital Machines” la propuesta de P. Rocchi, esta desarrollada en torno a la información en TIC (Tecnologías de la Información y Comunicación), con los conceptos de <Paradigma Digital> preferentemente, pero también asumiendo la parte imprescindible del <Paradigma Analógico>.

Tratados ambos para todo tipo de dispositivos, instrumentos y máquinas. Mediante las mismas, queda establecido en el procesamiento humano, primeramente desde el punto de vista del cerebro como sistema nervioso central, y previamente desde el sistema nervioso periférico, el que se conviva en interacción e interfase con los periféricos de las máquinas informáticas (aunque internamente al organismo animal y humano, los potenciales de acción, “spikes” cerebrales son netamente “digitales”, pero sin interacción natural con el exterior -excepto a través de instrumentos como la EEG, la MEG, la MRI funcional, etcétera).

Por esta parte, el tomar en nuestro caso, el sistema simbólico humano, el más importante, el lenguaje y su teoría semiótica, para desarrollar la teoría de la información, se convierte en central, porque incorpora el factor significado como imprescindible, a pesar de rozar con ello elementos de subjetividad que podrían distorsionar el enfoque científico cosa que P. Rocchi sortea y argumenta satisfactoriamente, en nuestra opinión (en la introducción de cabecera de este Apartado 2).

Su planteamiento de la información, nos pone en situación de abordar por otro lado, el próximo apartado, para avanzar hacia una mayor tipificación y una cuantificación del significado, incluyendo criterios para una teoría de la información ampliada (con el significado-semántico y su sentido “amplio”).

### **3. Teoría matemática del significado TMS (desde la teoría del <STC> relacionada – B. Moreno, 2017).**

El marco de la teoría de la información de P. Rocchi con el enfoque lingüístico y semiótico de partida, establece las bases para avanzar en lo que fueron planteamientos en la “Teoría matemática de la comunicación”, básicamente en el escrito que acompaña al de C. Shannon, por parte de W. Weaver en 1949.

Los planteamientos de W. Weaver (ver Apartado 1 pag. 3), quedan establecidos como una <contribución> a la teoría de Shannon, y son muy relevantes respecto a la incorporación del significado para la teoría de la comunicación. Aquí trataremos los tres primeros, (1ª) Los cambios en el <diagrama de bloques> (Figura 2), del de la comunicación de Shannon (Figura 1), (2ª) Estructura estadística a indagar del lenguaje, y (3ª) Variables canónicamente conjugadas de la “información” (compuesta de significante-léxico y significado-semántico, en nuestros términos).

Los planteamientos de las propuestas (4ª) y (5ª) se desarrollarán en un artículo en continuidad con el presente, referenciado como 5.2.

### 3.1 Cambios en el diagrama de bloques de C. Shannon (1948)

#### 31.1 El Receptor Semántico

La recepción del mensaje desde el punto de vista del <significado semántico>, supone varias cosas,

- a) Que el mensaje lleve codificación con una dimensión semántica
  - a1) Si el mensaje es una palabra, el código semántico podría ser una frase que la contenga con el significado a transmitir (de distintas opciones, como puede dar el significado: un diccionario, dando una definición implícita en una frase que contenga a esa palabra)
  - a2) Siguiendo con una palabra (denominada en lingüística “holofrase”), el código semántico podría ser una imagen o un sonido o ambas, que expresan el significado de la palabra en transmisión.
- b) Si lo anterior eran indicadores o indicios del <objeto-suceso> que nombra la palabra una situación diferente sería:
  - b1) Nombrar un sinónimo de esa palabra.
  - b2) Adjuntar un icono o ideograma (semejante a letra china, o emoticono)
  - b3) La función de adjuntar otros símbolos, como podrían ser: unas siglas, con acrónimos (por ejemplo UK, de United Kingdom), o ese término en otra lengua (traduciendolo, Reino Unido).
- c) También si se remite a un dispositivo más o menos tecnológico: un diccionario en forma de enciclopedia, un asistente de voz, una definición en Wikipedia, etcétera.

En base a las asociaciones conjugadas anteriores, el concepto de “entrelazamiento” (<ensamblaje> en el concepto y próximo de Rocchi), aquí es fundamental; es la interrelación entre el “significante léxico” y el “significado semántico”. Aparecen muchas variantes para establecerlo. Dos opciones desde aquí tratadas, adicionales y cuantificadas, como son por ejemplo:

- Si lo ciframos en coste de aprendizaje (de características analógicas)<sup>25</sup>.
- Si lo valoramos en términos de <información> de Shannon (de características digitales), tal como lo hemos referenciado con un bloque adicional del <Receptor Semántico> (en Figura 2, se verán detalles en punto 3.3).

En los dos casos el Receptor Semántico de W. Weaver, tiene dos contextos, uno analógico y otro digital.

Después de estudiar a P. Rocchi (2010), en el Apartado 2 más arriba, nos decantamos por el <paradigma digital>, si lo que queremos hacer es “comunicar”, pero lo haríamos con el <paradigma analógico> (entrada sensorial), si pretendemos que el receptor <aprenda> directamente, por lo cual es posible incluso, que en una visión de conjunto, se hiciera con ambos, comunicando y aprendiendo, dentro de sus funciones.

El Receptor Semántico quedaría con estos atributos, bastante completo.

#### 31.2 Fuente de Ruido Semántico

En el diagrama de bloques de la Figura 2, aparece en cajas remarcadas, y afectará a la parte semántica en exclusiva, y aparentemente, sólo en el proceso de comunicación.

---

<sup>25</sup> En la web [www.ingit.es](http://www.ingit.es) sección *En desarrollo*, sección de *Neurociencia y Psicología*, Artículo nº4

Es importante destacar que afectará al paquete semántico asociado (frase, imagen, sonido, sinónimo, etcétera), que en principio estarán con un nivel de ruido; los principios y teoría de Shannon referidos a este paquete exclusivo, responden similarmente al “Ruido de Ingeniería”. Pero el “entrelazamiento” habla de las reglas de interacción de la parte significativa con la parte de significado.

La mayoría de las palabras comunes tienen más de un significado, y hay pocos significados que dispongan de más de una palabra, o sea, los homónimos abundan, mientras que los sinónimos escasean.

No se sabe exactamente porque las lenguas son tan remisas con las palabras, y tan generosas con los significados. Es posible sea un principio de mínimo esfuerzo físico.

Todo ello configura un *potencial ruido semántico*.

¿Pueden estas reglas ser afectadas por un ruido distinto al de Shannon?

¿Una instrucción de algún programa informático, puede verse afectada por algún otro ruido?

¿La GGU (Gramática Generativa Universal de Chomsky), en un receptor humano, puede no ser infalible para el “entrelazamiento”, a pesar de ser básicamente biológica?, o aún más simple, ¿puede la sintaxis de una frase verse afectada por el ruido?

Darwin, decía de la sintaxis que es el fruto de <un órgano de extrema perfección y complejidad>. La organización de la gramática debería estar en la mente del niño en desarrollo, formando parte del mecanismo de aprendizaje del lenguaje, que le permite al niño dar sentido a los ruidos que escucha pronunciar a sus padres. La complejidad de la mente no es consecuencia de cómo se procesa el aprendizaje, más bien, el aprendizaje es consecuencia de la complejidad de la mente (Pinker, 1994).

Y ya ni hablamos del “entrelazamiento” de la palabra-mensaje con una imagen, un sonido, etc., aún más complicado de tipificar, aunque posible de realizar.

Las respuestas requieren investigación que aquí no se abordará.

### 3.2 La estructura estadística y algebraica del lenguaje, tanto de los Significantes, como de los Significados.

#### 32.1 Concepto como el de las cadenas de Markov y similares.

321.1 El contexto de una frase analizado únicamente el proceso combinatorio (con la propiedad de cadena de Markov), es insuficiente respecto al sistema formador de frases humano (mucho más selectivo).

No obstante, para el reconocimiento de voz, la cadena de probabilidades puede ayudar para asignación / confirmación de secuencias factibles. De hecho todo lo avanzado en tecnología de reconocimiento de voz, utiliza recursos matemáticos sacados de las probabilidades de uso de términos.

321.2 El proceso estocástico en cadena.

Requiere de la valoración de la probabilidad de que una palabra siga a otra. Otra variable del proceso, es la probabilidad de aparición según la longitud de la palabra, y muchas más estudiadas en lingüística, como la aplicación de la ley de Zipf.

## 32.2 Árboles sintácticos (S. Pinker, Bickerton etc).

### 322.1 Sintagmas y relaciones

El sintagma es una agrupación de palabras que funcionan como una unidad en una frase y que tiene normalmente un significado coherente. En gramática generativa se usa un conjunto de reglas que determinan la forma y significado de las palabras y las frases de una lengua en particular, tal y como se habla en una comunidad.

### 322.2 Reglas GGU (Gramática Generativa Universal) según N. Chomsky.

Es el diseño común y básico, a las gramáticas de todas las lenguas humanas. Su disciplina es la lingüística generativa. Subyace una única maquinaria encargada (físicamente ubicada en el cerebro en la <región perisilviana izquierda>, incluyendo a las áreas de Broca y Wernicke, cuya lesión produce afasias), y responsable de manipular símbolos, que sin excepción, cementa todas las lenguas del mundo. No se debe confundir con la estructura sintagmática superficial -s, y profunda -p, que son la categorías gramaticales de reglas que forman una oración, normalmente en forma arborescente también.

### 322.3 Estructuras fractales.

De recursividad como factor esencial (abundado por Hofstadter con textos como "Gödel, Echer, Bach", "Un extraño bucle" y "Analogía"). La recursión se utiliza para aplicar repetidamente sobre un grupo distinto de elementos, como pueden ser las palabras, que se puedan ir creando y analizando en un tamaño cualquiera. Por ejemplo ordenar palabras por "orden alfabético" siguiendo la secuencia de letras del alfabeto.

En las estructuras sintagmáticas, las palabras se agrupan en *sintagmas*, y estos se agrupan en otros *sintagmas mayores*. Normalmente se representan mediante un diagrama en *forma de árbol, que es intrínsecamente fractal*.

322.4 La especificidad estructurada de los espacios matemáticos: espacio algebraico ( $\xi_a$ ), espacio métrico ( $\xi_m$ ), y espacio vectorial-geométrico ( $\xi_v$ ), son ampliaciones sucesivas, matemáticamente equivalentes en unos casos (la aritmética de los números con la geometría) y en otros emergentes (como los efectos de nuevas dimensiones, por ejemplo de los números imaginarios de los reales).

#### 3224.1 Espacio algebraico ( $\xi_a$ )

Configura la "identidad" o "igualdad"; la equivalencia, la analogía, la <nitidez> de los términos.

$$A = A$$

$$A \neq B$$

Entidad,  $E \neq E^*$  (contiguo diferenciado; P. Rocchi, 2010)

NE (no E; P. Rocchi, 2010)

#### 3224.2 Espacio métrico ( $\xi_m$ )

Configura la definición de unidad; comparación y cuantificación.

Se plasman los números naturales, los enteros y la coordenada asociada.

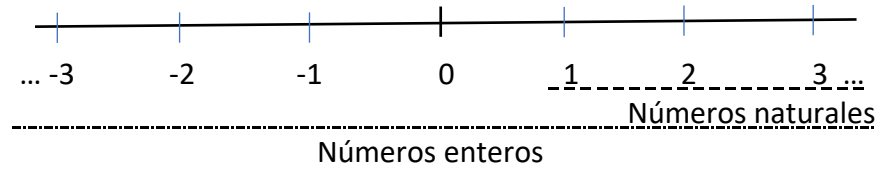


Figura 5

### 3224.3 Espacio vectorial-geométrico ( $\xi_v$ )

Es el espacio del plano cartesiano de partida, y el espacio geométrico como una aplicación (pueden ser representadas todas las figuras geométricas bidimensionales posibles).

Los números reales, imaginarios y complejos pueden ser representados en dos coordenadas ortogonales y asociadas (ordenada y abscisa).

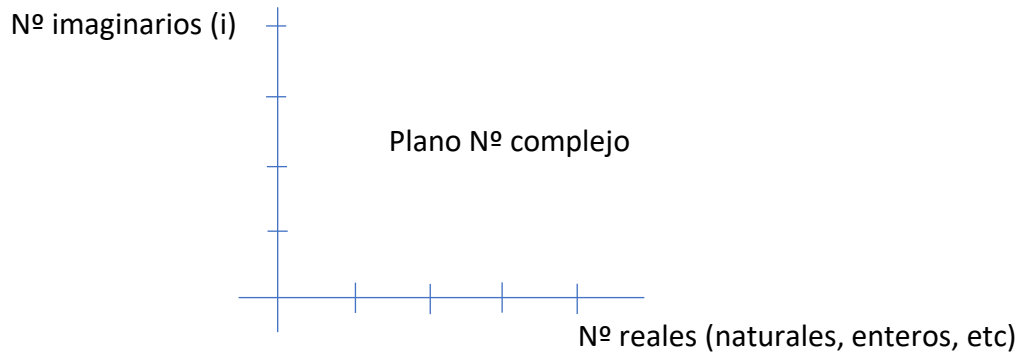


Figura 6

El modelo cartesiano, es probablemente el mejor modelo matemático para “conjuguar” (acepción de W. Weaver), el <significado> y el <significante> desde el modelo de P. Rocchi.

Tomando el triángulo semiótico visto en el Apartado 2, Figura 3, y proyectándolo sobre las coordenadas cartesianas en el espacio vectorial-geométrico ( $\xi_v$ ) de los números complejos, con los polos NE y E, en el plano

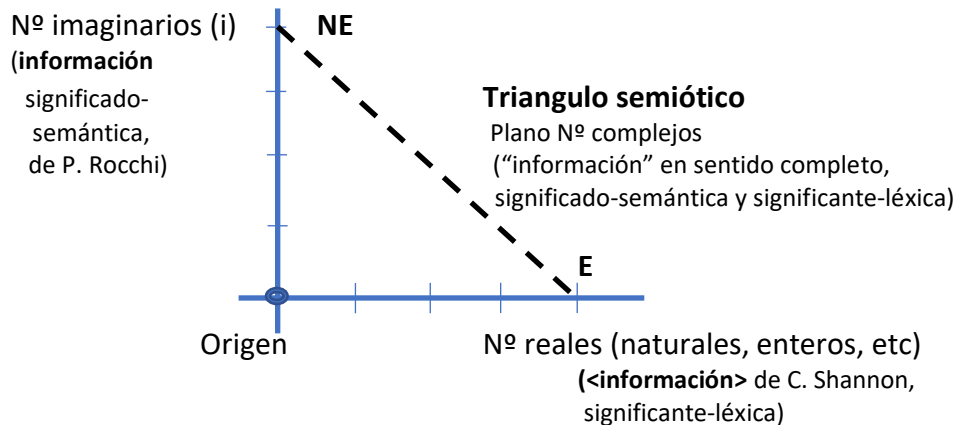


Figura 7

En el Origen de coordenadas, tenemos el “cero representación”, es decir está en un sentido el Objeto-Suceso (OS), motivo de la Referencia (R), y en otro sentido el observador, el dispositivo-instrumento-detector del OS, en sentido objetivo (un voltímetro, un termostato etc, indicador del objeto-suceso presente), que es acotable sin ser él, el mismo OS, aunque se codifique como entidad representada.

Convenimos que en el origen se da la equivalencia  $R \equiv OS$ , iniciándose en él las representaciones simbólicas E y NE, con carácter de entidad representacional codificada (objetivamente o convenida), y un acento distinto como: [*índices-indicios-indicadores*], [*iconos*] o [*símbolos*] de las entidades E y NE (Bearling 2009, Rocchi 2010, y otros).

Ver líneas de regresión en gráfico Fig. 8

### 3.3 La <información> de Shannon y el significado, “canónicamente conjugados” (en la propuesta de W. Weaver, 1949)

Nuestro objetivo aquí, siguiendo esa propuesta, y partiendo de la gráfica de aprendizaje para tipificar la “irrupción lingüística” en bebés-niños (B. Moreno, 2020; aquí en nota 24 y gráfico en Anexo 1), es el adoptar en el código del lenguaje natural lo que vincule “significante” con “significado”, y como hemos visto reflejado en el espacio algebraico  $\xi_a$ , métrico  $\xi_m$ , y vectorial  $\xi_v$ . Pero veamos como se estructura lingüísticamente en una tabla de equivalencia de términos:

NIVEL	AUTOR	MARCO	
		SIGNIFICANTE	SIGNIFICADO
1	Saussure Peirce Morris		
2	P. Rocchi (2010)	Significante Semiótico <E>	Significado Semiótico <NE>
3	C. Shannon (1948) W. Weaver (1949)	<Información>	Significado
4	Propuesta 1 (nuestra, con propósito desambiguante)	<b>Significante-léxico</b>	<b>Significado-semántico</b>
5	Propuesta 2	Palabra aislada	Palabra en frase; contextualizada
6	Hinzen – Poeppel (2011) ( con un sentido semejante, Putnam (1985) Bloom (1985), etcétera)	Significante con < <b>significado estrecho</b> >	Significado en frase, < <b>significado amplio</b> >
7	Propuesta 3: “ <b>irrupción verbal</b> ”, en el desarrollo bebé-niño (ver Anexo 1)	<u>Número de palabras de vocabulario</u> manejadas sin contexto, aisladas.	<u>Número de palabras en una frase</u> , articuladas en contexto.

Tabla 1

El nivel 7 con la propuesta 3, es la base que tomaremos para la evaluación cuantificada de la información con significado. Se aplican como indicadores en función de su procedencia, la palabra aislada y la palabra contextualizada en una frase.

La “irrupción verbal” en niños a los 18 meses, surge con significado estrecho con las aproximadamente 75 palabras de vocabulario, aprendidas y pronunciadas. También pueden concebirse como frase de una sola palabra u “holofrase”, que tienen un significado amplio muy bajo, de hecho referenciado como <estrecho>. En este contexto, los morfemas son las unidades mínimas dotadas de *significado*, en los que se pueden descomponer las palabras (*in- /micro- /onda- /bili- /-dad*), dando significados que se pueden componer.

**Resumen de datos tabulados en el desarrollo de un bebe-niño en meses de edad:**  
(ln: logaritmo natural o neperiano, para analizar regresión)

### SIGNIFICANTES

nº palabras	ln - Nº de palabras	Edad (meses)
0	0	13
<b>75</b>	4,32	<b>18</b>
375	5,93	24
1050	6,96	36

Tabla 2

### SIGNIFICADOS

nº palabras	ln – Nº de palabras	Edad (meses)
0	0	15
1	0	18
2	0,69	21
11	2,40	36
18	2,89	48

Tabla 3

#### 33.1 <Entrelazamiento>, o relación de interacción entre significante-léxico ST (E), y el significado-semántico SD (NE).

Vamos a definir a continuación, la <información> en términos de Shannon, y tanto de <E>, como de <NE>. En el ámbito lingüístico, E y NE, son el número de palabras sueltas y el número de palabras combinadas en una frase, cuya magnitud acotan el número de palabras que la componen.

Estas dos magnitudes las llamamos  $H_E$  y  $H_{NE}$ ; respectivamente información aportada por el número de símbolos significantes, por un niño en desarrollo cifrado en meses, y el número de símbolos significados del número de palabras combinables en una frase en edades sucesivas.

#### Curvas de regresión entre E y NE, muestran su independencia matemática como variables

Edad (meses)	ln ST (significante)	ln SD (significado)
13	0	
15		0
18	4,32	0
21		0,69

La independencia es condición imprescindible para la atribución dimensional referida en las Figuras 6 y 7, y que aquí se verifica, con la **convergencia de la regresión** de ambas.

24	5,93	
36	6,96	2,4
48		2,89

Tabla 4

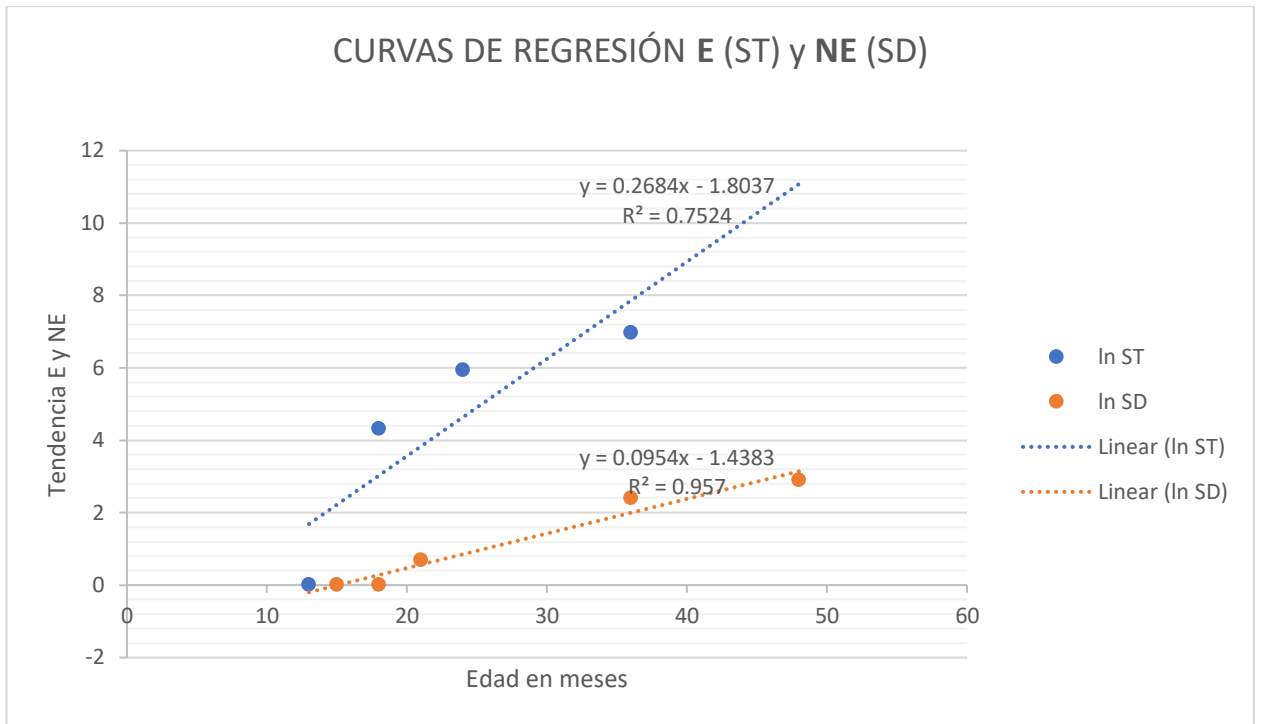


Figura 8

Aproximadamente hacia los 24 meses, llegan los niños a disponer de unas 375 palabras, y con frases compuestas por 3 palabras en cada frase en las que inicia su “composición” (monos adiestrados, enlazan máximo unas decenas de 2 términos o palabras de comprensión en cada frase, muchas de ellas gestuales).

El entrelazamiento NE – E, como información  $H_E$  y  $H_{NE}$ , para una teoría de la información ampliada con el <significado amplio>,

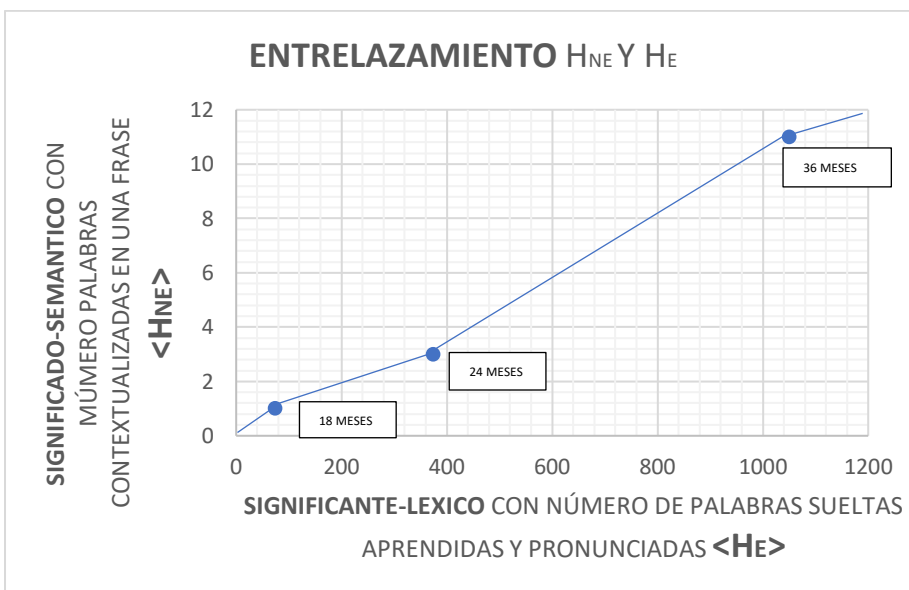


Figura 9



Cada palabra es un símbolo que va en el mensaje de comunicación en el modelo de Shannon, por tanto la <información> de Shannon para  $H_E$  : 375 palabras, y simplificando considerando todas las palabras equiprobables,

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

En el caso del significado, su potencial informativo nos lo da una cota máxima de  $H_{NE}$  : 3 palabras, pero que son combinadas de las 375 palabras significantes E.

Establecemos por tanto a un nivel combinatorio, cuantas frases pueden crearse sin restricciones con:

$$\begin{aligned} n &= 375 \text{ palabras a combinar} \\ r &= 3 \text{ palabras combinadas en cada frase} \end{aligned}$$

Las combinaciones son entonces:

$$C_n^r = \binom{n}{r} = \frac{n!}{r!(n-r)!} \approx 8.859.500 \text{ frases de 3 palabras}$$

Por tanto la información con significado contextualizado en frases de 3 palabras, siendo simplificada equiprobables:

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

Si vamos a nuestro espacio algebraico  $\xi_a$ , métrico  $\xi_m$ , y vectorial-geométrico  $\xi_v$ , para representarlo, definimos la “información conjugada” o “entrelazada”, como **HH**.

Para un objeto-suceso  $OS_n$  dado, de cada una de las “n” palabras, que en este caso es el vocabulario significativo de 375 palabras, tenemos un nivel contextualizado en frases de 3 palabras.

El posible significado de estas 3 palabras en un niño con un nivel de desarrollo de 24 meses,

$$\mathbf{HH}^{24}_{375/3} = \mathbf{H}^{24}_{375} + i \mathbf{H}^{24}_{375/3} \quad \text{en el plano adoptado de } n^{\circ} \text{ complejos}$$

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

### 33.2 Extendemos el entrelazamiento como < sistema conjugado > (W. Weaver), al mismo fenómeno de desarrollo en los bebés-niños de 36 meses.

La fracción no lineal, pasando de 375 palabras a los 24 meses, a las 1050 palabras de vocabulario (significante-léxico) a los 36 meses, consideradas con distinto número de palabras, y estableciendo también los términos de significado-semántico, en el que se llegan a crear frases de 11 palabras en esos 36 meses.

Siguiendo las pautas de 33.1 pero con distinto número de palabras, y considerando también su no-linealidad,

n = 1050 palabras a combinar  
r = 11 palabras en cada frase

Podemos calcular la <información> de Shannon, que es:

$$H_E = \log_2 1050 = \underline{10,04 \text{ bits}}$$

Y en una información con significado amplio, las combinaciones son entonces:

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

La información con significado contextualizado en frases de 11 palabras para un formato simplificado equiprobable (como en el “protolenguaje”, en las descripciones de Bickerton & Calvin, *Lingua exmachina*, p.53) para cada palabra,

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

La información conjugada o entrelazada HH en el espacio matemático ( $\xi_a, \xi_m, \xi_v$ ), queda,

$$HH_{1050/11}^{36} = H_{1050}^{36} + i H_{1050/11}^{36} \quad \text{en el plano de } n^\circ \text{ complejos}$$

$$HH_{1050/11}^{36} = \mathbf{10 + i 95}$$

Cuya generalización es la misma obviamente, y que queda reflejada en las fórmulas:

Generalizando:  $HH_{OS_n}^R = H_E^R + i H_{NE}^R$  [4]

(ecuación en el plano de los  $n^\circ$  complejos)

Siendo, **R**: Referencia, para un Observador, Dispositivo-instrumental-detector-actuador, según índice-indicio-indicador / icono / símbolo (Rocchi, 2010), de un OS dado “n”

**OS<sub>n</sub>**: Objeto-suceso “n”, origen de las representaciones simbólicas “n”

**E**: Entidad simbólica lingüísticamente <significante> (Rocchi, 2010), del OS

**NE**: Entidad simbólica lingüísticamente <significado> (Rocchi, 2010), del OS

**i**: unidad imaginaria del plano complejo, en el espacio vectorial-geométrico  $\xi_v$ .

La fórmula desplegada en función de las probabilidades de uso de cada representación simbólica “n”,

$$HH_{OS_n}^R = - \sum_1^{nR} p_E \log_2 p_E - i \sum_1^{nR} p_{nNE} \log_2 p_{nNE} \quad [5]$$

### 33.3 Apreciaciones para el contenido concreto del mensaje.

Con lo tratado en 33.1 y 33.2, y hemos evaluado la información canalizada durante la irrupción verbal del lenguaje a los 24 y a los 36 meses.

Tomando a los 24 meses,  $HH_{375/3}^{24} = 8,55 + i 23,1$  (en bits)

Podemos adoptar un contenido concreto de la frase de 3 palabras, por ejemplo, con la siguiente frase:

#### **“Papá está contento”**

Dicha por el bebe-niño de 24 meses a su madre, y que una vez formulada (para tener el significado explícito que se quiere trasladar a la madre), puede cuantificarse.

Para ello se envía a través del proceso de comunicación, figurado por el “diagrama de bloques” (Figura 2), la frase léxica de 8,55 bits de información de Shannon y de 23,1 bits de información con significado.

Todo sucede dentro de un código sintáctico formado por : nombre, verbo y complemento. Se envía con una capacidad de permutaciones de 3 palabras concretas,  $3! = 6$ . La información asociada a estas 6 combinaciones, supone un valor informacional de **2,6 bits**, que frente a los 8,55 representa un **30%** de ocupación de la comunicación potencial del canal en la fracción léxica. La semántica contiene otros factores de posible definición, veamos algunas posibilidades, balanceadas desde los 23,1 bits máximos del haz semántico posible.

Puede llevar las siguientes asociaciones léxicas:

- “Papá”: “*progenitor masculino*”, en el que ni “progenitor”, ni “masculino”, pertenecen al universo verbal de las 375 palabras que en este estado de desarrollo tiene a los 24 meses, y que maneja el niño. Para la composición de la frase, tienen que escogerse varias palabras del diccionario de las 375, o atribuciones directas que asocian “papá”, con esa persona que está en casa.
- “está”, presente del *verbo “estar”*, estado factual de esa persona.
- “contento”, “*estado emocional de satisfacción*”, también por atribución directa posiblemente.
  - Otra opción de recepción es asociar uno o varios modo/s - sensorial/es, específicos:
    - Por ejemplo, foto del padre sonriendo
    - Por ejemplo, grabación audio del padre con carcajadas
  - Una opción híbrida de formato abstracto (lingüístico) – concreto (sensorial): mezclando modos verbales y sensoriales.

Finalmente se hará balance del <mensaje específico>, en base a su universo léxico (de 375 palabras en el nivel de desarrollo con 24 meses) y los términos de los diccionarios que adjuntemos (léxicos con mas palabras; o bien asociando mensajes sensoriales en modos visual, auditivo, etc).

Si por cada palabra del mensaje, la explicáramos como mínimo por otras 2, y como máximo hasta las 375 restantes, la resultante es menos 3 de la frase explícita, estamos hablando de  $375 - 3 = 372$  términos conocidos a elegir para los dos términos adicionales.

El “pequeño mundo”<sup>26</sup> de este diccionario, es bastante restringido, pero es el que tiene ahora el niño.

No obstante hay que pensar que también se recurre al “cúmulo” o “clusters” de memoria, que en forma de engramas tienen los modos sensoriales inscritos, no léxicos, que tienen su propio diccionario (visual, auditivo, etc) y que pueden hibridarse con el del léxico cuando son evocados. Estos últimos no van por el canal oral que el niño está usando con su madre, pero no siendo presentes en la comunicación, si lo son en la interpretación de información que el niño realiza al lanzar la observación hablada a su madre, básicamente cumpliendo las funciones de Receptor Semántico.

Lo que no es léxico o semántico, no entra en esta valoración, ceñida a los términos lingüísticos.

## CONCLUSIÓN

Las ecuaciones [4] y [5], describen y cuantifican, sin entrar en consideraciones de redundancia, seguridad o aspectos complementarios de interpretación, la información significativa-léxica y la del significado-semántica, que permiten determinar y completar el **valor informacional del significado** que nos ocupa, en el segundo término de dichas ecuaciones.

Este valor es objetivamente, y tomando la “irrupción verbal” como referencia a los 36 meses, de 95 bits / 10 bits = **9,5 veces mayor** que la información de Shannon, reflejando la <información> léxica con *significado estrecho* en las palabras aisladas, permitiendo así una cuantificación del **significado amplio** por la contextualización del significado en la frase.

Por lo tanto el significado amplio es prácticamente **1 orden de magnitud mayor** que el significado estrecho, y por tanto a igualdad de capacidad simbólica y aprendizaje, tiene 10 veces más de necesidad de tiempo (como recurso de aprendizaje) para la <accesibilidad cognitiva>; o por otro lado, otros recursos para adquirir el significado, que en el caso del vocabulario léxico son más livianos, en su acepción de significado estrecho.

**El significado**, para el caso lingüístico, y durante el periodo de aprendizaje de vocabulario y de la utilización del habla de un niño, queda razonablemente acotado en su valoración de **información** en comunicación lingüística (la incorporación de medios audiovisuales, por ejemplo, permite una descripción paralela que se puede analizar, pero que aquí no es el caso).

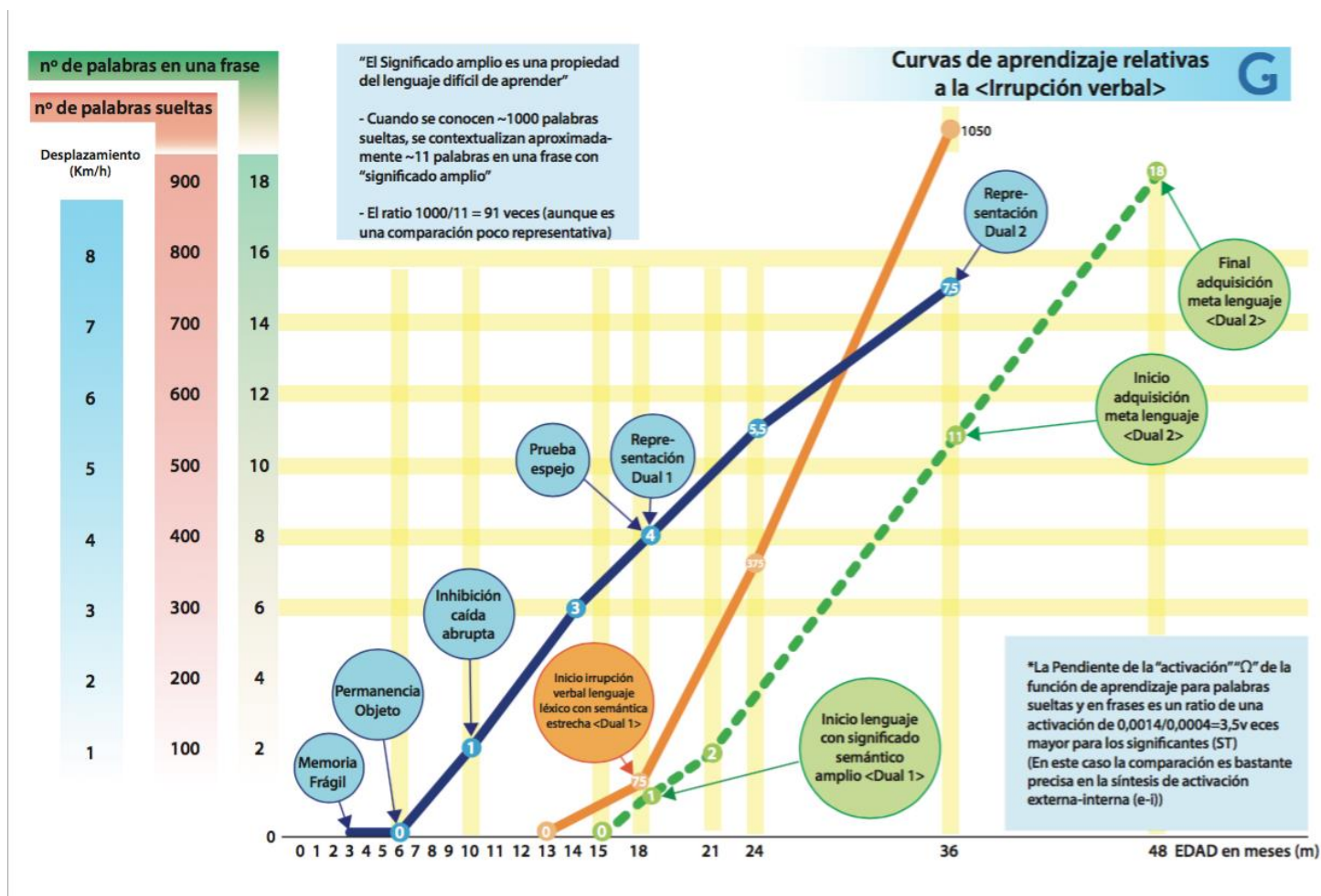
**La dimensión adicional** que supone, representada en el plano complejo (con numeración compleja), significa un incremento de capacidad del canal de comunicación, para acceder a los significados, en un niño de 3 años. Esta capacidad adicional, que es de un orden de magnitud mayor que el vocabulario (x 10 aproximadamente), y que se despliega en los meses de desarrollo que van de los 13 a los 48 meses (curvas naranja y verde en el Anexo 1), permiten ver su potencial explicativo, prácticamente completamente desplegado.

También el Receptor Semántico del diagrama de bloques de la Figura 2, representa a esta **nueva dimensión** en el modelo de la información y comunicación de C. Shannon, y que ya W. Weaver supo avanzar, la del significado.

---

<sup>26</sup> Solé, R. (2009). *Redes complejas. Del genoma a Internet*. Tusquets Ed., Metatemas.

## Anexo 1



[www.ingit.es](http://www.ingit.es), en la pestaña *En desarrollo*, apartado *Neurociencia y Psicología*, Artículo 4

- Desde estas gráficas se deduce la gráfica de la Figura 9, concretamente combinando la curva verde y la naranja bajo el parámetro de los meses de edad media que precisa el niño para su desarrollo.

### Nota

Se agradece encarecidamente a Jordi Sos, Ingeniero en Telecomunicaciones y amante de la Calidad como criterio de excelencia, su contribución en la búsqueda bibliográfica continua, a veces de petición reiterativa por mi parte, y siempre atendida. Gracias.