

Complejidad e Ingeniería

El rol de los lenguajes en nuestro desempeño

Engineering and complexity

The role of languages in our performance

Gerardo Febres¹

¹ Universidad Simón Bolívar, Venezuela, *gerardofebres@usb.ve*

Resumen

Se enfocan los lenguajes como vehículos para transmitir la descripción de sistemas. En ese sentido un lenguaje se manifiesta como el modelo de un sistema pero también como el andamiaje que sostiene nuestros pensamientos y que nos permite estructurar nuestra interpretación de los sistemas. Algunos experimentos sencillos se muestran para ilustrar la formación de lenguajes a partir de reglas sobre patrones de símbolos. Se presentan ideas para la evaluación de la efectividad de lenguajes y modelos sencillos de su evolución. Resultados de trabajos de investigación sobre los lenguajes literario y musical se ofrecen como evidencia de que esta visión de los lenguajes tiene validez. Finalmente se sugiere una relación entre Complejidad, Lenguaje e Inteligencia, y se resalta su importancia en el desempeño del científico y del ingeniero.

Palabras clave: lenguajes, descripción, sistemas complejos, entropía, escala de observación.

Abstract

Languages are focused as vehicles to convey the description of systems. In that sense a language manifests itself as the model of a system but also as the scaffolding that sustains our thoughts and allows us to structure our interpretation of systems. Some simple experiments are shown to illustrate the formation of languages from rules on symbol patterns. Ideas for evaluating the effectiveness of languages and simple models of their evolution are presented. Results of research work on literary and musical languages are offered as evidence that this view of languages is valid. Finally, a relation between Complexity, Language and Intelligence is suggested, and its importance in the performance of the scientist and the engineer is emphasized.

Keywords: languages, complex systems, descriptions, entropy, observation scale.

1. Introducción

Hay un vínculo entre razonamiento y lenguaje.
Pensamos haciendo uso de algún lenguaje.

Podemos asociar unos fenómenos con otros y establecer la relevancia de los efectos de unos agentes con otros mediante una especie de computador basado en uno o varios lenguajes que utilizamos para soportar nuestros modos de

cálculo mental: nuestros modelos de la naturaleza. En este trabajo se persigue como objetivo el hacernos conscientes del impacto que los lenguajes tienen sobre la calidad de nuestro desempeño como científicos e ingenieros. Para ello se presentan algunos de los mecanismos que conduce cada lenguaje a su proceso evolutivo y se presenta a los lenguajes y al raciocinio como dos manifestaciones de un mismo fenómeno, la inteligencia.

Estas conjeturas son sometidas a pequeños experimentos con lenguajes numéricos para así dar fuerza a su validez. Adicionalmente se citan resultados encontrados en trabajos previos que nos llevan a concluir sobre la evolución conjunta de los lenguajes y la inteligencia. Finalmente se muestra que es posible utilizar expresiones de lenguajes para identificar su origen.

2. Un modelo evolutivo de los lenguajes

Los lenguajes están constituidos por patrones de símbolos y conjuntos de reglas capaces de transmitir información con una cierta efectividad. La conveniencia de transmitir información de un modo cada vez más efectivo debe producir el uso más frecuente de los símbolos más compactos y de las reglas que conducen a mayor precisión en las ideas transmitidas, y que al mismo tiempo son más fácilmente entendidas por los receptores del mensaje. Desde los años 1940's Flesch [1], [2] en sus trabajos sobre *lecturabilidad* de textos en idioma inglés estableció escalas para cuantificar el impacto de la longitud de las frases y las palabras en la comprensión del mensaje. Cuando un símbolo puede acortarse sin comprometer el significado con otro símbolo preexistente, seguramente adoptará la forma

más compacta. De igual manera, las reglas, que en los lenguajes naturales vienen dadas por la posición de palabras de la clase cerrada (preposiciones, artículos, terminaciones para la formación de adverbios, terminaciones para la conjugación de verbos, etc.) se modifican progresivamente para alcanzar una uniformidad razonable y funcionar con la precisión que caracteriza a un lenguaje poderoso.

En lo referente a la cantidad de símbolos con que cuenta el lenguaje, es sencillo suponer que con el tiempo nuevos símbolos aparecen para referirse a conceptos emergentes o para referirse con más detalle a aspectos específicos de viejos conceptos. En los lenguajes naturales esto ocurre con la aparición de nuevas palabras, generalmente de la clase de las palabras abiertas (nombres, adjetivos, etc.).

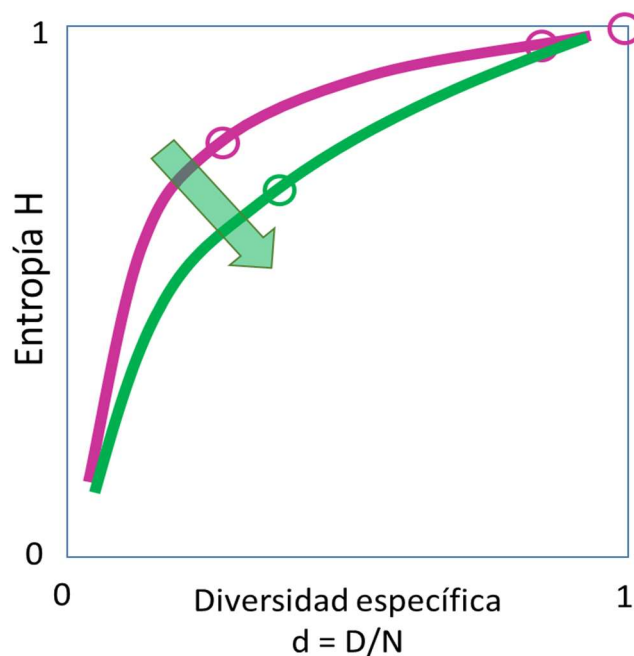


Figura 1 Esquema evolución de los lenguajes en el plano entropía – diversidad específica.

Por otra parte, cuando el ajuste en el uso de una regla produce resultados convenientes,

seguramente vendrá acompañado de una reducción en la entropía asociada al texto. Este ajuste entonces, por su conveniencia, tendrá más probabilidades de *sobrevivir* y quedar como parte del lenguaje en una condición cuasi-estable.

Los efectos combinados en el crecimiento de la diversidad específica, y la reducción en la entropía [3] causados por la mejor auto-organización de las reglas y la posición relativa de los símbolos, hacen suponer que los lenguajes evolucionan como se ilustra en la Fig. 1, La descripción de un sistema, un discurso, o simplemente una idea expresados con un lenguaje poco evolucionado, queda representado sobre un lugar de relativa alta entropía y baja diversidad específica. En la medida que el lenguaje evoluciona debe esperarse que los puntos que representan la misma idea se desplacen hacia zonas de menor entropía y mayor diversidad específica.

3. Modelo simbólico de los lenguajes

De las formas de percepción, la más prominente en los humanos es la vista. Podemos describir con gran precisión y detalle lo que vemos. Y podemos captar el aspecto de algo, incluso cuando está a gran distancia. Nuestro lenguaje natural se ha desarrollado para expresar las percepciones visuales en un grado mayor que con otros sentidos. Para contrastar, considérense los sentidos del gusto y el olfato. Si se nos pide describir un sabor o un olor, rápidamente nos encontraremos en la necesidad de referirnos a alguna sustancia cuyo olor o sabor sea parecida. Pero difícilmente podremos describir en términos abstractos un olor. No tenemos un lenguaje desarrollado para describir olores y debemos

basarnos en comparaciones para ubicar un olor en algún contexto.

Así, las entidades que existen en dimensiones espaciales difíciles de representar para los humanos, pueden ser concebidas a través de su representación análoga como formas o colores que sí somos capaces de interpretar. Una idea que hemos utilizado es la de dibujar la intensidad relativa de los símbolos que componen la expresión de un sistema. Ordenando los símbolos de acuerdo a su frecuencia, se obtiene como resultado una línea curvada que resulta característica de la entidad descrita. Esta fue la observación que hizo Zipf [4] cuando notó que para el idioma inglés la frecuencia de cada palabra distinta obedece a una distribución potencial según la expresión

$$f_r = \frac{f_1}{r^g}, \quad (1)$$

donde r es el lugar que ocupa un símbolo de acuerdo al número de apariciones, f_r es el número de veces que aparece el símbolo r y g es un valor real que para el inglés es cercano a 1.

Podemos ver el hallazgo de Zipf como una forma de describir al idioma inglés. También es posible aplicar este método a otras lenguas naturales y establecer comparaciones interesantes entre ellas. Efectivamente es un tema de intenso estudio que cae fuera del alcance de este trabajo. Más allá de la aplicación directa de este método de representación de un texto, se puede utilizar este tipo de gráfico (perfiles) como método para asociar a una forma visible, la descripción de una entidad realizada en términos que a primera vista, no podemos interpretar.

La Figura 2 muestra la forma que, de acuerdo a esta representación, tiene cuatro discursos representados en idioma español.

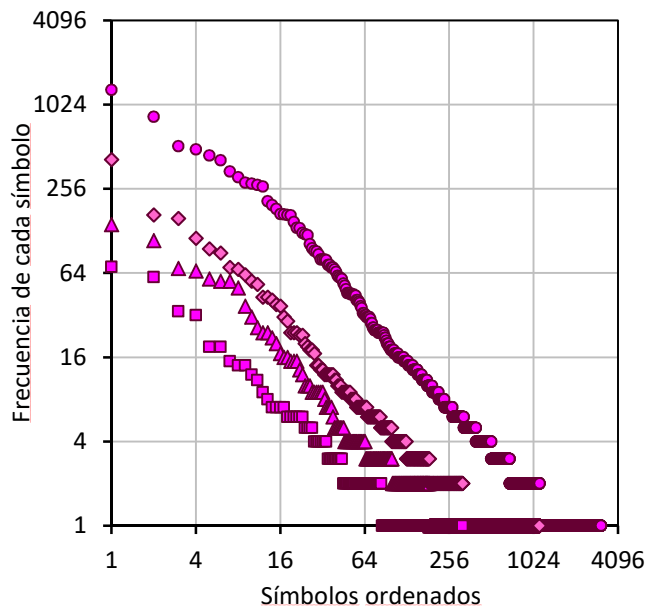


Figura 2 Cuatro perfiles de discursos expresados en español. Tomado de [5].

SQUARE 1936. DOLORES BARRURI
 TRIANGLE - 1982. GABRIEL GARCÍA MÁRQUEZ
 RHOMBUS - JOSESARAMAGO. VALENCIA,
 CIRCLE - CAMILO JOSE CELA. LACOLMENA. CAP 1

Los perfiles de los cuatro discursos representados en la Figura 4 se aproximan a una distribución lineal sobre ejes con escalas logarítmicas. Tal como lo expresa la Ecuación (1). Además exhiben pendientes con valores muy cercanos entre sí, lo que sugiere que se trata de diferentes manifestaciones del mismo lenguaje.

La representación de perfiles, como los mostrados en la Figura 2, es posible una vez que los símbolos que componen una descripción son reconocidos. En el caso de la versión escrita de los lenguajes naturales es muy sencillo reconocer palabras, que a su vez son consideradas como símbolos. Sin embargo,

en la mayoría de los casos, el registro escrito de las manifestaciones de un sistema, no hay palabras ni símbolos fáciles de reconocer. Se hace necesario aplicar Principios para identificar el tamaño y la ubicación de los símbolos relevantes de la descripción.

4. La Escala fundamental

La Escala Fundamental es el conjunto de símbolos más relevante para la interpretación de una descripción. Teniendo en cuenta que la máxima capacidad de transmisión de información ocurre cuando es mínima la entropía asociada a la interpretación, se ha construido un algoritmo [6] capaz de evaluar diferentes formas de observar el mismo texto, y seleccionar de ellas, aquella que conduce a la mínima entropía. La Figura 3 muestra un esquema muy básico del algoritmo Fundamental Scale Algorithm (FSA).

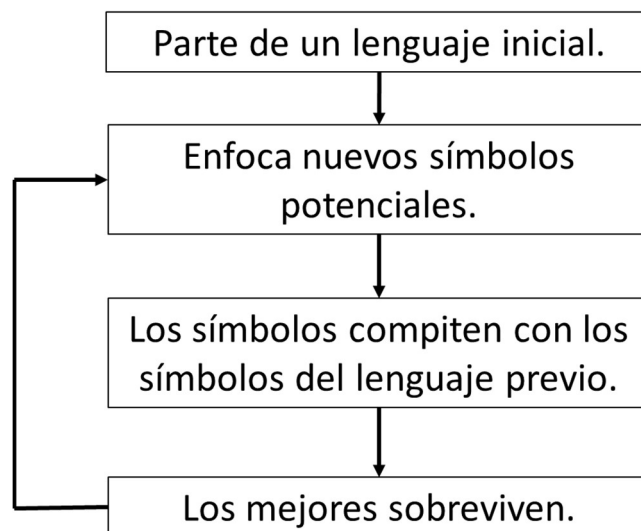


Figura 3 Esquema operativo del algoritmo de la Escala Fundamental FSA.

La aplicación del FSA ha permitido la construcción de perfiles asociados al sonido de piezas musicales. Varios cientos de piezas

musicales de varios géneros y estilos han sido estudiados con este método [7]. Algunos de los perfiles resultantes se muestran en las Figuras 6 y 7.

5. Aplicaciones y resultados

La Figura 4 muestra los valores de entropía según la diversidad específica de discursos escritos en español e inglés y códigos de programas de computadora. Se observa una clara separación de los tres lenguajes evidenciando que los lenguajes artificiales no han tenido tiempo suficiente para evolucionar hasta los niveles de baja entropía exhibidos por su contraparte de los lenguajes naturales.

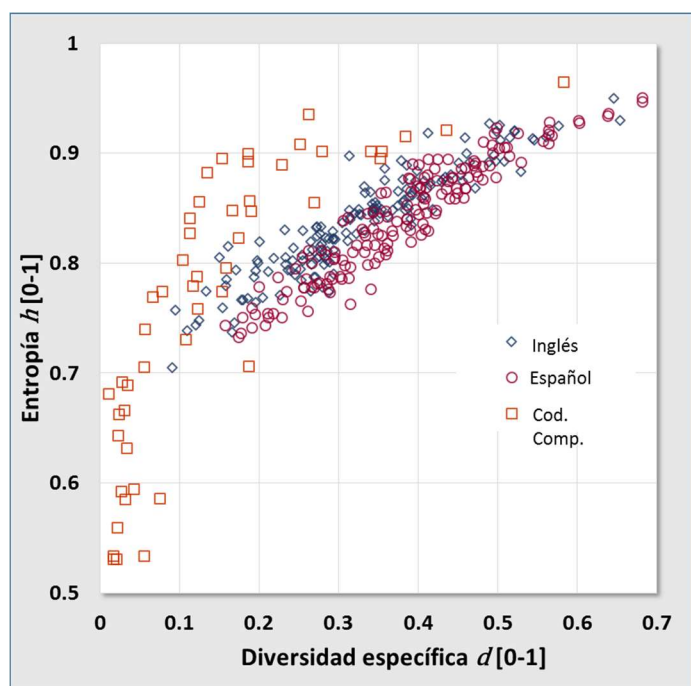


Figura 4 Comparación entre lenguajes naturales (inglés y español) y lenguajes artificiales (códigos de programación). Obtenido de [5].

La Figura 5 muestra los resultados de comparar los estilos de redacción entre escritores y escritores que ganaron el Premio Nobel de

Literatura. La comparación se realizó para discursos escritos en español e inglés. Para ambos lenguajes se incluyeron discursos originalmente escritos en idiomas diferentes y llevados hasta la expresión evaluada mediante traducciones del texto original. No se observan alteraciones importantes debidas a posibles efectos de las traducciones.

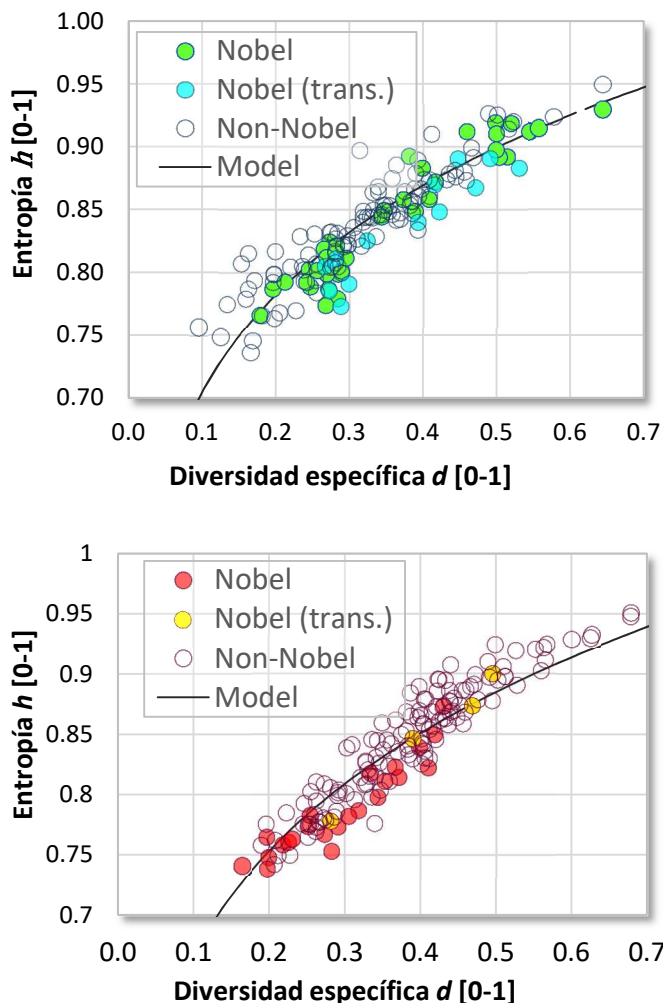


Figura 5 Comparación entre estilos de redacción de escritores y merecedores del Premio Noble de Literatura. Gráfico superior muestra discursos en inglés y el inferior en español. Obtenido de [8].

La Figura 6 muestra perfiles representativos de la música asocia a dos periodos: Barroco y Romántico. Puede decirse que estos

perfiles representan una especie de línea media para todas las piezas musicales incluidas de cada periodo.

La diferencia en la forma evidencia que el método es capaz de asociar formas distintas a expresiones sonoras distintas. Con lo cual es posible caracterizar estas expresiones musicales.

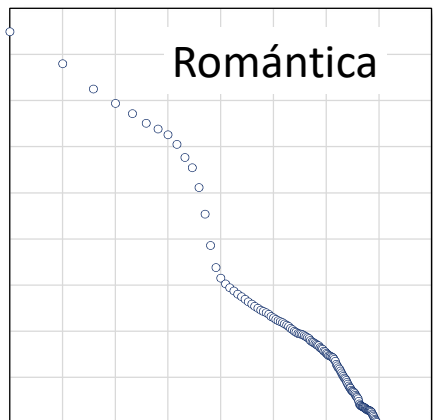
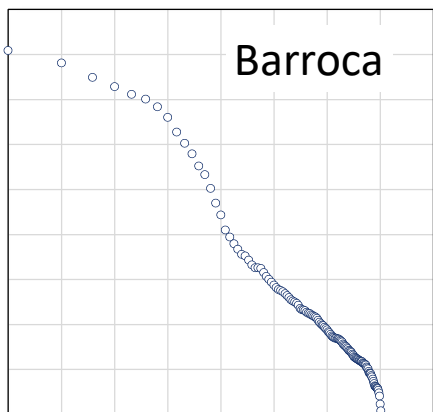


Figura 6. Dos perfiles representativos de géneros musicales correspondientes a los periodos barroco y romántico. Ambos ejes se muestran a escala logarítmica. Las unidades no son relevantes, lo importante es la forma. Tomado de [7].

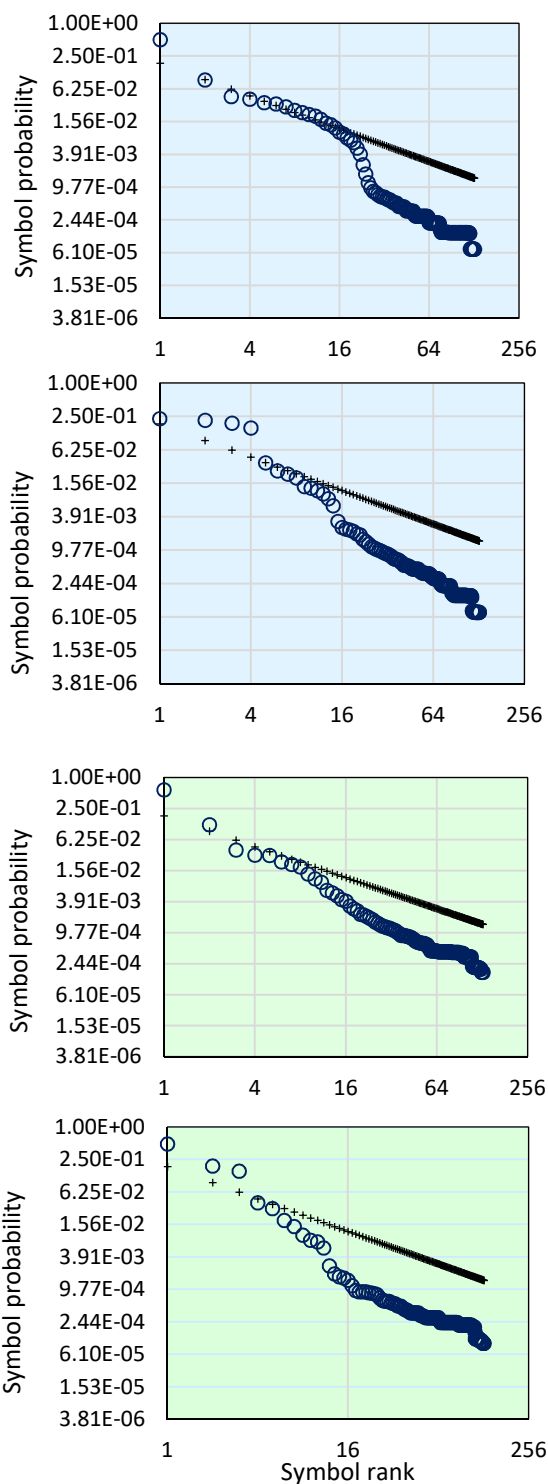


Figura 7. Perfiles asociados a pares de interpretaciones de la misma pieza musical. Arriba: interpretaciones en piano y órgano de Toccata y Fugga de JS Bach. Abajo: interpretaciones de dos arreglos distintos de El Diablo Suelto. Tomado de [7].

La representación de descripciones abstractas, inicialmente ininteligibles, a través de la forma que adoptan los perfiles de símbolos asociados, ha servido para comparar géneros musicales. Pero estas capacidades de reconocimiento del método van más allá de los géneros y estilos musicales. La Figura 7 muestra los perfiles asociadas a dos piezas musicales, cada una interpretada con distintos instrumentos y arreglos. Para el humano es fácil identificar la identidad de cada pieza. Sin embargo los sonidos son radicalmente diferentes lo que hace suponer que los perfiles resultantes muestren diferencias igualmente radicales. No es el caso. Los perfiles mostrados en la Figura 7 ilustran como para la Toccata y Fuga de J.S. Bach y para El Diablo Suelto, conocida pieza de la música tradicional venezolana, los perfiles exhiben formas afines de manera correspondiente que pueden ser utilizadas para que el computador ‘escuche’ e identifique una pieza musical.

6. Discusión

La comparación de lenguajes naturales artificiales y naturales ilustrada en la Figura 2 sirve como elemento de prueba de la veracidad del método evolutivo de los lenguajes. Evidentemente los lenguajes artificiales, representados por códigos de computadora, no han tenido tiempo ni uso suficiente para competir en capacidad con sus contrapartes naturales. Y efectivamente, los niveles de entropía y diversidad específica mostrados en la Figura 2 son consistentes con estas consideraciones, comprobando así la veracidad de nuestros supuestos y el modelo evolutivo que hemos adoptado.

La relevancia del plano entropía–diversidad específica queda de relieve al poder reconocerse en este plano diferencias más pequeñas como las que existen entre el español y el inglés. Más aún, diferencias tan sutiles como las que separan los estilos de redacción de los escritores con los que han merecido el Premio Nobel de Literatura, son claramente detectadas y evidenciadas en la Figura 5. La agrupación de los discursos de estos últimos en la región de baja entropía y alta diversidad específica, nos impide resistir la tentación de decir que el uso que ellos hacen del inglés y el español tiende a ser más ‘evolucionado’ que el de los demás escritores. Muestra contundente de la efectividad de estos métodos.

Este trabajo parte de algunos supuestos para proponer métodos para convertir percepciones ininteligibles, en formas gráficas cuantificables y comparables. Estos métodos se ofrecen como una alternativa interesante para ‘ver’ a través de un gráfico lo que originalmente no era perceptible por la vista.

7. Conclusiones

El modelo simbólico de lenguaje, propuesto originalmente por Zipf y la entropía como propiedad para evaluar la capacidad de un lenguaje para transmitir información, sirven como base para la comparación cuantitativa entre medios para la descripción de sistemas complejos. El concepto de Escala Fundamental extiende la aplicación de estos métodos a descripciones totalmente abstractas de un sistema como son aquellas realizadas en lenguajes que caen en el rango desconocido para el observador.

La concepción de escala de observación como parámetro para la interpretación de la descripción de un sistema complejo, ofrece ventajas para la selección de la interpretación del sistema y el modelado de su comportamiento, empoderando el lenguaje mismo y extendiendo nuestra propia capacidad de comprensión. La ciencia y la ingeniería evolucionan al ritmo que lo hace el dominio de los lenguajes con los que construimos campos de estudio.

Referencias

- [1] R. Flesch, *The art of readable writing*. New York: Harper & Brothers, 1949.
- [2] R. Flesch, *The art of clear thinking*. Harper & Brothers, Barnes and Nobel Books.
- [3] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.
- [4] G. K. Zipf, *Human Behavior and the principle of least effort: An introduction to human ecology*. New York: Addison-Welesly, 1949.
- [5] G. Febres, K. Jaffe, and C. Gershenson, "Complexity measurement of natural and artificial languages," *Complexity*, vol. 20, no. 6, pp. 429–453, 2015.
- [6] G. Febres and K. Jaffe, "A Fundamental Scale of Descriptions for Analyzing Information Content of Communication Systems," *Entropy*, vol. 17, pp. 1606–1633, 2015.
- [7] G. Febres and K. Jaffe, "Music viewed by its Entropy content: A novel window for comparative analysis," *arxiv.org*, vol. arxiv:1510, 2016.
- [8] G. Febres and K. Jaffe, "Quantifying structure differences in literature using symbolic diversity and entropy criteria," *J. Quant. Linguist.*, vol. 24, no. 1, 2017.