



## **La Ley de Zipf y los historiadores.**

Luis Eduardo Cortés Riera.

[cronistadecarora@gmail.com](mailto:cronistadecarora@gmail.com)

Esta curiosa Ley fue descubierta desde la ciencia de la lingüística por el estadounidense George Kinsley Zipf (Illinois, 1902- Massachusetts, 1950), cuando realizaba trabajos de filología comparada en la década de 1930. Empleando la estadística encontró que las palabras más utilizadas parecían estar gobernadas por unos patrones de aparición. A modo de ejemplo, la palabra más utilizada que ocupa el primer puesto sería el doble de veces más utilizada que la segunda más utilizada, el triple de veces que la tercera, el cuádruple de veces que la cuarta, y así sucesivamente. Esta Ley parece cumplirse en todas las lenguas, incluidas las construidas artificialmente como el Esperanto.

Zipf trabajó en la Universidad de Harvard y se interesó en el mandarín chino. Sus descubrimientos sobre el lenguaje, el mayor vehículo de creación de cultura, le han llevado también a ser uno de los autores más citados en el área de la informática y la documentación científica, especialmente en el área de recuperación científica de información y la indización automática.

Zipf no estaba solo en el empeño, pues otros investigadores venían en el camino con semejantes propuestas: el criptógrafo estadounidense Claude Elwood Shannon, descifrador de códigos secretos durante la Segunda Guerra, creador de la teoría de la información, y el polaco Benoit Mandelbrot: las nubes no son caóticas, tienen un orden.

En textos clásicos como *El origen de las especies*, publicado en 1859 por Charles Darwin parece que se cumple indefectiblemente la Ley de Zipf. Zipf estudió este fenómeno con textos en inglés y, por lo visto, se cumple. Si cogemos la versión original de *El origen de las especies de Charles Darwin* (1859) vemos que la palabra más utilizada en el primer capítulo es "the", con una aparición de cerca de 1.050, mientras que la segunda es "and", apareciendo cerca de 400 veces, y la tercera es "to," apareciendo unas 300. Aunque no de forma exacta, se puede ver que la segunda palabra aparece la mitad de veces que la primera y la tercera un tercio.

La frecuencia de rangos en el uso de las palabras parece ser un factor independiente de la cultura, dirá Noam Chomsky. Zipf se interesa particularmente de la gigantesca novela de James Joyce de 1922: *Ulises*, manuscritos en latín, en alemán, chino.

Pero lo más desconcertante es que esta misteriosa ley se puede trasladar al mundo natural y al mundo de lo social. Eventos azarosos como el Tour de Francia y sus ganadores por país, la frecuencia del número de tornados y huracanes en un día en los Estados Unidos, el número de artistas en función del precio medio de su obra, la frecuencia de guerras y el consiguiente número de fallecidos, el código genético del ADN, la frecuencia de los terremotos que varía con su magnitud, el crecimiento poblacional de las ciudades: Nueva York tiene el doble de habitantes que Chicago, que tiene el doble de Filadelfia, la diversidad de las especies según el área de su hábitat, el diámetro de los cráteres de la Luna, en economía la distribución de riqueza. Una misteriosa progresión matemática que no obedece a una causa universal.

Nuestro personaje, profesor de la lengua alemana, cometió un grave error en 1941, durante la Segunda Guerra Mundial: intentó aplicar sus notables descubrimientos a la geopolítica, al avance de las tropas nazis en Europa. Justificó en un libro la expansión hitleriana argumentando que obedecía a urgencias cuantificables que hacían que Alemania no pudiera evitar invadir a sus vecinos.

Actualmente sabemos gracias a Zipf que las regularidades estadísticas en el lenguaje son reales y profundas. Los modelos de lenguaje funcionan precisamente porque esas regularidades existen y son explotables. Si el lenguaje fuera totalmente impredecible ChatGPT no existiría.

Las 10 páginas más visitadas en la actualidad son Google (60, 49 millones de visitas cada mes) Youtube (24,32 millones), Facebook (19, 98 millones), Baidu (9,77 millones), Wikipedia (4, 69 millones), Twitter (3,92 millones), Yahoo (3,74 millones), Pornhub (3,36 millones), Instagram (3,21 millones) y Xvideos (3,19 millones). Viendo estas cifras, se puede ver que Google es el doble de visitado que Youtube, el triple que Facebook, más del cuádruple que Baidu.

### **La Ley Zipf y los historiadores.**

A nuestro parecer resulta muy útil que los cultivadores de la ciencia de Clío conozcamos esta curiosa ley de Zipf. Ella tiene puntos de acercamiento con la campana de Gauss que nos enseñaron en la universidad y que es muy útil para los historiadores de la economía o de la demografía histórica. Aplicada a los escritos del Libertador Simón Bolívar daría resultados que supongo revelarían muchos aspectos desconocidos del Padre de la Patria venezolana.

En nuestra experiencia fue muy útil para detectar la frecuencia de las adscripciones a las cofradías y hermandades coloniales y republicanas en la ciudad de Carora, Venezuela, y con ello medir estadísticamente el fervor y el entusiasmo religioso durante largos cuatro siglos en la “Ciudad levítica de Venezuela.” La hipótesis de trabajo, o esta fecunda imprudencia de método, dice Michel Vovelle, de afirmar que entre los gestos de la práctica y la fidelidad religiosa hay una correlación tosca, por cierto, pero positiva e indiscutible. Se trata de una Tesis Doctoral que contiene 76 tablas estadísticas, diez gráficos de líneas y barras, gráficos circulares. No utilizamos ninguna fórmula, pues nos valimos de una estadística meramente descriptiva.

La Ley de Zipf, para sorpresa nuestra, se cumple con bastante aproximación en la cofradía del Santísimo Sacramento, los hermanos fundadores siempre duplican a los que “entran” como hermanos en el segundo mes tras la fundación de la cofradía, y los hermanos que se inscriben el tercer mes son la mitad de los que se inscriben el segundo mes de actividad de la hermandad.

En el futuro pienso establecer cuáles son las palabras más empleadas en mi Tesis Doctoral. Me asesoraré con expertos en computación y daré uso de las máquinas de última generación computacional, así como la Inteligencia Artificial. Sería muy interesante saber cuál es la palabra que más veces salió de mis manos, así como la que menos usé.

**Fuentes consultadas:**

Cortés Riera, Luis Eduardo. (2003). Iglesia católica, cofradías y mentalidad religiosa en Carora, Venezuela. Editorial Académica Española. Leipzig, Alemania.

Font-Clos, F., Boleda, G. y Corral, Á. (2013) A scaling law beyond Zipf's law and its relation to Heaps' law. *New Journal of Physics*, 15. doi.org/10.1088/1367-2630/15/9/093033.

Montemurro, M. A. (2001). Beyond the Zipf–Mandelbrot law in quantitative linguistics. *Physica A: Statistical Mechanics and its Applications* 300: 567 - 578.

. Nahum Montagud Rubio. (2020, enero 9). Ley de Zipf: un curioso fenómeno social y matemático. Portal Psicología y Mente. <https://psicologiymente.com/cultura/ley-zipf>

Vovelle, Michel (1985) Ideologías y mentalidades. Editorial Ariel S. A. Barcelona,

Zipf, George Kingsley (1932): *Estudios seleccionados sobre el principio de frecuencia relativa en el lenguaje*. Cambridge, Massachusetts. Estados Unidos.

Carora,

Estado Lara,

República Bolivariana de Venezuela,

31 de mayo de 2026.