



La historia del conocimiento como un relato del lenguaje

Knowledge history as language tale

Iwan Mascolo

ORCID: 0009-0007-2910-5323

Centro de Filosofía para la Investigación Stanislaw Strba (CENFISS).
mascoloiwang@gmail.com

Recibido: octubre 2025

Aprobado: enero 2026

Resumen

Este ensayo explora la evolución del lenguaje como matriz epistemológica fundamental y como sistema generativo autónomo. Desde las tradiciones orales míticas hasta las ambiciones formales del *Ars Magna* de Llull y la *Characteristica universalis* de Leibniz, el texto rastrea cómo el lenguaje ha moldeado (y ha sido moldeado por) la búsqueda del conocimiento universal. Basándose en teorías estructuralistas y semióticas (Saussure, Peirce), sostiene que el significado surge no de la referencia sino de la diferencia relacional y la mediación interpretativa. El ensayo culmina con un análisis del procesamiento del lenguaje natural (PNL) contemporáneo, donde el lenguaje se vectoriza y modela estadísticamente, permitiendo incluso como términos inventados o sin sentido adquieran un significado funcional. Este cambio revela un nuevo paradigma: el lenguaje ya no refleja simplemente el pensamiento: construye la realidad, incluso en ausencia de un sujeto consciente. Las implicaciones abarcan epistemología, ontología, ética y semiosis artificial, desafiándonos a repensar los límites entre sentido, estructura y autonomía en la era del lenguaje computacional.

Palabras clave: Lenguaje, conocimiento, símbolo, semiótica, vector

Abstract

This essay explores the evolution of language as both a foundational epistemological matrix and an autonomous generative system. From mythic oral traditions to the formal ambitions of Llull's *Ars Magna* and Leibniz's *characteristica universalis*, the text traces how language has shaped—and been shaped by—the pursuit of universal knowledge. Drawing on structuralist and semiotic theories (Saussure, Peirce), it argues that meaning arises not from reference but from relational difference and interpretive mediation. The essay culminates in an analysis of contemporary Natural Language Processing (NLP), where language is vectorized and modeled statistically, allowing even invented or nonsensical terms to acquire functional meaning. This shift reveals a new paradigm: language no longer merely reflects thought—it constructs reality, even in the absence of a conscious subject. The implications span epistemology, ontology, ethics, and artificial semiosis, challenging us to rethink the boundaries between sense, structure, and autonomy in the age of computational language.

Keywords: Language, knowledge, symbol, semiotics, vector

Introducción

La historia del conocimiento humano está íntimamente ligada a la historia del lenguaje. Desde las primeras manifestaciones simbólicas en cavernas hasta los sistemas formales de la lógica contemporánea, el lenguaje ha sido mucho más que un medio de expresión: ha sido el terreno donde se ha cultivado la conciencia, la herramienta con la que se ha tallado el pensamiento, y el espejo donde la humanidad ha intentado reflejar la totalidad del mundo.

Antes de la existencia de la ciencia como disciplina, antes incluso de que la filosofía se emancipara del mito, el ser humano ya intuía que nombrar era crear. En las cosmogonías antiguas, el verbo tiene poder generativo: decir es hacer, y el mundo se ordena según las palabras nombradas. Esta intuición, atraviesa culturas y épocas, revelando una verdad profunda: el lenguaje no solo describe la realidad, la constituye.

Este ensayo se propone explorar esa dimensión fundacional del lenguaje, no como simple vehículo de ideas, sino como matriz epistemológica. A través de un recorrido desde la oralidad mítica hasta

los intentos modernos de construir lenguajes universales, se examinará cómo el lenguaje ha sido el suelo fértil donde germinan las categorías del saber. Se analizará su papel en la estructuración del pensamiento filosófico, en la formalización del conocimiento científico, y en la configuración de la subjetividad moderna.

Más allá de su función comunicativa, el lenguaje será abordado aquí como una tecnología simbólica permitiendo la abstracción, la inferencia y la sistematización. En este sentido, se argumentará como todo conocimiento es, en última instancia, conocimiento lingüístico: incluso las matemáticas, aspirando a la pureza formal, no escapan a la mediación de signos y convenciones.

La tesis central que guía esta investigación puede formularse así: el lenguaje no es un accesorio del pensamiento, sino su condición de posibilidad. Esta afirmación, que puede parecer radical, será desarrollada a lo largo del texto mediante el análisis de momentos clave en la historia de la epistemología, desde la filosofía griega hasta los proyectos modernos de codificación total del saber.

El ensayo que sigue abordará el surgimiento del lenguaje como fundamento epistemológico, explorando cómo los primeros filósofos lo utilizaron para delimitar lo verdadero, lo racional y lo universal. Se mostrará cómo el paso del mito al logos no fue solo una transformación narrativa, sino una revolución ontológica que redefinió la relación entre palabra y mundo.

1. La génesis del conocimiento

Desde los albores de la civilización, el lenguaje ha sido mucho más que una herramienta de comunicación: ha sido el medio por el cual el ser humano ha estructurado su pensamiento, ha dado forma a su mundo y ha delimitado los contornos de lo cognoscible. En este sentido, el lenguaje no solo transmite conocimiento, sino que lo constituye. Esta afirmación, lejos de ser una mera provocación filosófica, encuentra sustento en la evolución histórica del pensamiento occidental.

En las primeras etapas del pensamiento humano, el mito funcionaba como una forma de explicación totalizante. Las narraciones míticas no distinguían entre lo simbólico y lo empírico; eran relatos que condensaban experiencia, creencia y orden social. Sin embargo, con el surgimiento de la filosofía griega, se produce una ruptura epistemológica: el logos comienza

a desplazar al mythos como forma legítima de saber.

Este tránsito no fue inmediato ni pacífico. Los sofistas, por ejemplo, comprendieron el poder del lenguaje como instrumento de persuasión y manipulación, mientras que Sócrates y Platón lo reivindicaron como vía para alcanzar verdades universales. La mayéutica socrática, basada en el diálogo, revela que el conocimiento no se impone, sino que se extrae del alma mediante el uso preciso del lenguaje.

Aristóteles lleva esta concepción aún más lejos al formalizar el silogismo como mecanismo lógico. En su obra, el lenguaje se convierte en arquitectura del pensamiento: cada proposición, cada término, cada relación sintáctica refleja una estructura ontológica. Así, el lenguaje no solo describe el mundo, sino que lo ordena y lo hace inteligible.

Este giro epistemológico inaugura una tensión que persiste hasta nuestros días: ¿puede el lenguaje ser un medio neutro para representar la realidad, o está irremediablemente cargado de presupuestos ontológicos y culturales? La ciencia moderna, al buscar un lenguaje ideal —matemático, formal, universal— intenta resolver esta paradoja, pero en el proceso revela otra: cuanto más se depura el lenguaje, más se aleja de la experiencia humana concreta.

La tesis que subyace a este capítulo puede resumirse en una reformulación del dictum de Protágoras: el lenguaje es la medida de todas las cosas. No porque todo sea relativo, sino porque todo lo que puede ser pensado, dicho o conocido está mediado por estructuras lingüísticas. Esta idea prepara el terreno para el siguiente capítulo, donde se explorará el sueño moderno de una “Característica Universal”: un lenguaje perfecto que permita representar todo conocimiento sin ambigüedad ni contradicción.

2. La característica universal (De Lull a Leibniz)

Y descendió el Señor a ver la ciudad y la torre que edificaban los hijos de Adán, y dijo: He aquí, el pueblo es uno solo, y todos tienen la misma lengua; y han empezado esta fábrica, ni desistirán de sus ideas hasta llevarlas a cabo. Ea, pues, descendamos, y confundamos allí mismo su lengua, de manera que el uno no entienda el habla del otro. Y de esta suerte les esparció el Señor desde aquel lugar por todas las tierras, y cesaron de edificar la ciudad, de donde se le dio el nombre de Babel o Confusión (Génesis, 11, 5 – 8).

Pensar en el lenguaje como una abstracción de carácter general suele parecer antinatural, dada la existencia de tantos dialectos distintos, cada uno con su propia estructura única proveniente de siglos de evolución. Y aunque la pregunta de qué viene primero —el lenguaje o la cultura— evoca la misma sensación del dilema del huevo y la gallina, cierto es que ambos, cultura y lenguaje, son dos caras de la misma moneda: la visión del mundo y de nosotros como grupo.

Distintas lenguas equivalen a distintas cosmovisiones, religiones, morales, éticas; pero, independientemente de todas estas disyunciones, todos los grupos tienen algo en común, siendo esta la misma causa de su desentendimiento con otras civilizaciones: el habla. Pues el ser humano es el animal que posee lenguaje (*zoon logon echon*). Siendo la lengua tanto el origen del problema como su única solución posible, ¿acaso no sería posible unir en conversación a distintas culturas con una única habla, una des confusión del mundo?

La confección de una lengua capaz de crear puentes entre hablantes de distintos dialectos, sin importar lo disonantes que estos fueran, fue la gran ambición de Ramón Llull (c. 1232–1316), quien, motivado por la idea de unir en diálogo interreligioso a las distintas culturas conocidas del medievo, con el fin de demostrar de forma cuasi aristotélica las verdades de la fe, no tomándolas de autores anteriores, sino creándolas uno mismo en un “Arte Mayor” (*Ars Magna*).

La idea era simple: ya que cada religión tiene sus propios dogmas e interpretaciones únicas de sus textos sagrados, busquemos demostrar —no a través de estos textos, sino mediante la razón— las verdades de lo divino. Y para esto es necesario un encuentro entre las dos razones máximas concebidas hasta entonces por el hombre: lo platónico y lo aristotélico, lo geométrico y lo lógico en pro de la teología.

Aunque la manifestación fue evolucionando a lo largo de la vida de Llull, el concepto fundamental se mantuvo igual hasta el fin de sus días: un sistema basado en polígonos concéntricos —por lo general, círculos— divididos en sectores que representaban conceptos, atributos o manifestaciones de lo divino. Las figuras se giraban hacia un sector X, permitiendo que las propias proporciones de los polígonos señalaran el resto de sectores o atributos, los cuales, finalmente combinados, generaban un silogismo cuyo origen estaba en las proporciones de los objetos geométricos y no en interpretaciones humanas, resolviendo de alguna forma el problema de las premisas lógicas de Aristóteles (ver introducción).

Esta *Ars Magna* es el primer sistema del que se tenga registro alguno donde el lenguaje construye significados —o, en este caso, silogismos— sin necesidad de una interpretación humana activa, sino más bien a través de las propias reglas de dicho sistema. En este caso, las proporciones geométricas en el plano son las que dictan cómo se generan las proposiciones; de forma similar a como la aguja de una brújula que apunta al norte siempre da la cara contraria al sur, sin importar que el usuario esté de acuerdo o no.

Puede que, en su momento histórico, Llull no dimensionara la escala del acercamiento que estaba teniendo hacia el lenguaje, adelantándose siglos y siglos a la visión del lenguaje y el habla que se tiene en la actualidad: sin significados ni significantes, donde el habla solo necesita del hablante al inicio y al final, ya que la construcción activa nace del propio lenguaje como estructura separada de los sujetos.

A lo largo del Medievo, la *Ars Magna* de Llull no proliferó como él hubiera querido, y los estudios sobre el lenguaje como estructura lógica continuaron muy apegados a la visión de Aristóteles, de la mano de autores como Guillermo de Ockham y Pedro Abelardo. Aunque la idea del lenguaje como una maquinaria capaz de construir verdades universales a través de nada más que de sí mismo seguiría resonando

hasta alcanzar a su siguiente portador en la modernidad, ya bien alejada de la época escolástica. Si bien se podría decir que Llull dio el primer paso hacia una visión sistemática del lenguaje, no sería hasta la llegada de Gottfried Leibniz (1646–1716) que este sueño alcanzaría una forma mucho más moderna y reconocible a nuestros ojos: la “Lengua característica universal” (*lingua characteristica universalis*) como un lenguaje simbólico universal, sin ambigüedades ni imprecisiones, fue la ambición de este racionalista; un álgebra del habla, si se quiere.

Así como Descartes con la epistemología, y más tarde Hobbes con la política, Leibniz buscaba matematizar el lenguaje, permitiendo manipular las premisas involucradas a través de permutaciones, conjunciones, sustracciones, etc. Sin limitarse al verbo aristotélico, el lenguaje natural, en toda su ambigüedad, debía transformarse en algo exacto, con reglas estrictas pero que permitieran un comportamiento dinámico para adaptarse a toda la complejidad que representa el conocimiento en relación con la razón y la experiencia.

A pesar de ser un concepto fascinante, el proyecto de Leibniz con esta lengua matemáticamente universal era —si cabe— incluso más ambicioso que el de Llull con su *Arte Combinatorio*, ya que el fin del alemán era dar solución a todas las riñas existentes entre los distintos campos del saber y dentro de sí mismos. Esto supondría no solo una revolución en la epistemología, sino también una utopía donde no habría debate sobre ningún tópico relacionado al conocimiento en sí mismo. Pero no hace falta más que un vistazo a la historia para percatarse de que esta lengua universal no parece haber cumplido su propósito en lo absoluto.

Leibniz fue un genio en muchas áreas —prácticamente en todas las que abordó—, pero la lengua universal simplemente se quedó en el campo de la especulación. Debido a la suma complejidad de abarcar todo lo humano en una reducción sin pérdida de información en el proceso —algo perfectamente factible en computación y matemáticas—, fue imposible adaptarlo al lenguaje natural. Por lo tanto, fue posiblemente —y con el perdón de Newton— la mente más prolífica del Renacimiento y, aun así, los fragmentos de este trabajo de Leibniz pisaron fuerte en la filosofía y lingüística de sus sucesores.

Aunque inconclusa, la empresa de una lengua matemática trajo consigo un abanico epistemológico inmenso para lo que sería la lingüística moderna y la lógica, siendo esta idea no solo la antesala de la lógica booleana y descriptiva, sino también el primer contacto formal del lenguaje con el álgebra y la aritmética. De allí surgiría el origen de conceptos vitales para el nacimiento de la lingüística como ciencia.

Con el paso de los siglos, el estudio de la lengua como objeto en sí mismo quedó relegado a un segundo plano. El principal enfoque de la lingüística se centró entonces en la historicidad de los distintos dialectos del mundo conocido: entender las raíces del español en el latín, del inglés y el alemán en un germánico antiguo. Mientras tanto, la pregunta por qué es el lenguaje quedaba relegada a una respuesta cómoda: “Un espejo del pensamiento”. Así lo creyeron muchos, incluso autores como Wilhelm von Humboldt (1767–1835). Pero fue justamente uno de los conceptos tratados por Leibniz el que abrió las puertas al renacer de la lingüística y al estudio del lenguaje como ente en sí.

Para Leibniz, era menester encontrar una forma de reducir todos los elementos pertenecientes al lenguaje a estados primarios, del mismo modo que los números pueden descomponerse en sus factores primos. Así, el lenguaje puede ser estudiado desde sus raíces estructurales, de las cuales surge toda la complejidad adyacente a la experiencia humana de la lengua —misma que detuvo las cavilaciones de Leibniz sobre la lengua universal.

Si tomamos la lengua como un simple espejo de la realidad enmarcado en nuestra mente, la aparición de conceptos tan icónicos como problemáticos en teoría del conocimiento, como lo son: El infinito, el tiempo, el ser y demás abstracción; se convierten en auténticos quebraderos de cabeza para todo aquel que como Leibniz busque una verdad universal coherente con la realidad en el lenguaje, y también para quien como Humboldt, dicte que el lenguaje es un mero espejo del pensamiento, pues nadie nunca ha visto al “ser” y haciendo uso de un razonamiento cartesiano: ¿Cómo puede un ser finito tener una idea de lo infinito, si no ha experimentado jamás lo infinito ni tiene los medios para construirlo a partir de lo

finito?

Dada la existencia de estos imposibles del lenguaje como representación de la realidad, solo queda preguntarse como surgen estos cisnes negros de la realidad, que a nuestro entender son tan comunes y corrientes como cualquier otro concepto concebible por la lengua, y la respuesta parece ser más que obvia: Si existe un objeto finito, debe existir lo infinito, si existe una premisa P por puro razonamiento debe existir su negado ¿No es así?

Lo anterior fue solo razonamiento, pero he aquí el detalle; para que el negado de P exista, se debe tener ya conocimiento del P positivo, por lo que un elemento halla su existencia en su partición con otro: Lo finito y lo infinito, lo bello y lo horrible, lo cocido y lo crudo. Las palabras encuentran sentido frente otras palabras, el lenguaje se genera a si mismo pues igual a la Ars Magna escolástica, el lenguaje natural se nutre de su propia estructura interna sin necesidad de intervención del sujeto, pasando de un racionalismo a estructuralismo.

3. Estructuras y símbolos: De Saussure a Pierce

A lo largo de este escrito ya se han mencionado múltiples veces conceptos como “significados”, “significantes”, etc., pero el momento de desglosarlos a profundidad no llega sino hasta ahora; a mediados del siglo XIX e inicios del XX, con Ferdinand de Saussure y el nacimiento de la lingüística moderna.

“La lengua es un sistema de diferencias sin términos positivos”. La anterior cita del Curso de lingüística general de Saussure engloba bastante bien la idea sobre el lenguaje que se ha estado construyendo a lo largo del ensayo: la lengua como un sistema que se da sentido a sí misma y no solo como una herramienta para dar sentido a la realidad. Esa es la máxima de la lingüística según Saussure, que le otorga al lenguaje no solo la cualidad de ser un espejo de la mente —como se había planteado anteriormente—, sino también la capacidad de producir sus propias dinámicas, elementos y relaciones únicamente a través de la estructura intrínseca del propio lenguaje.

Para Saussure, el lenguaje está compuesto por una unidad elemental que llamó signo lingüístico, cuya naturaleza comprende dos manifestaciones: el significado (la imagen mental) y el significante (la imagen acústica); siendo estos, respectivamente, la representación conceptual y la forma sonora de esa representación en el lenguaje. Perro, como animal mamífero, sería el significado; y la palabra perro, como sustantivo común singular, sería el significante. Estas dos manifestaciones del signo lingüístico son la base del lenguaje como sistema y, por ende, también de todas las lenguas habidas y por haber. Por lo tanto, podría decirse que este signo lingüístico es un símil —si se quiere— del elemento primario del lenguaje que Leibniz buscaba para su Característica Universal.

Hasta este punto, la construcción del signo lingüístico no dista mucho de la visión más espiritual o psíquica de Wilhelm von Humboldt —salvando las diferencias ontológicas que conlleva—, pero esto se debe a que esta explicación es solo parcial. El signo lingüístico no se limita a representaciones léxicas. ¿Qué pasa con las proposiciones, artículos y conjunciones? Las conjunciones, por ejemplo, solo representan causalidad. Y haciendo alarde de empirismo: ¿alguien alguna vez ha presenciado la causalidad como una realidad tangible? Si la respuesta a esta pregunta es negativa, ¿cómo podemos explicar la aparición de estas palabras —estos significantes sin significado aparente— que no tienen equivalente en la realidad?

Esta cuestión es resuelta por Saussure con la idea nuclear del estructuralismo lingüístico —y tal vez la más importante de toda su obra—: “En la lengua no hay más que diferencias sin términos positivos”.

Con esta afirmación, Saussure deja sentada la idea de que la creación de significado va más allá de una mera cuestión de representación: incluye una capa contextual en la construcción de sentido para el significante. Básicamente, las palabras pueden adquirir su significado a través de su contexto o posición dentro de la estructura gramatical, y por su relación con otros significantes. De esta forma, se resuelven

las incógnitas planteadas al final del capítulo anterior y se abre la puerta a una forma completamente nueva de entender la lengua como un ente en sí mismo, separado de los sujetos.

Esta visión del lenguaje da una respuesta elegante y natural a los “cisnes negros” del lenguaje presentados anteriormente. Por ejemplo, el concepto de lo infinito surge como una deducción basada en otro concepto —en este caso, lo finito—: si lo finito existe, lo infinito también podría hacerlo. Y aunque no lo haga, la conceptualización de lo infinito ya forma parte del lenguaje: un significante sin un significado real en el mundo, pero perfectamente válido en la estructura lingüística.

Esta respuesta de Saussure, aunque válida, aún presenta un área de ambigüedad que solo puede ser tratada de forma adecuada analizando el lenguaje no solo como sistema, sino también en su relación con el conocimiento.

Haciendo un pequeño repaso: en la Grecia clásica de Platón y Aristóteles imperaban dos grandes metodologías para alcanzar el conocimiento válido, siendo estas la mayéutica (para Platón) y la lógica (para Aristóteles). La primera busca develar un conocimiento que se encuentra implícito en el sujeto a través del diálogo, casi como un recorrido por distintos conceptos y razonamientos hasta llegar a una conclusión satisfactoria. La segunda, en cambio, se encuadra en la creación de razonamientos basados en proporciones lógicas, a través de las propias leyes de la lógica creadas por Aristóteles, inspirado en la sofística presocrática.

Teniendo ambos modelos en mente, es posible hacer una conexión interesante con la teoría saussureana del lenguaje, ya que este último divide el lenguaje en dos manifestaciones humanas: el habla (*parole*) y la lengua (*langue*); siendo la primera la manifestación individual y concreta del lenguaje, encausada únicamente por los hablantes, y la segunda, el sistema de significados y significantes que ya se expuso anteriormente. A ambos conceptos se les puede vincular, respectivamente, con la mayéutica (para el habla) y la lógica (para la lengua): dos formas de generar conocimiento. La primera, a través de la pluralidad, tomando el lenguaje como un juego con más de un jugador —todos los involucrados en el proceso de diálogo—; y la segunda, mediante una exploración individual del lenguaje como un juego entre el sujeto y la estructura del lenguaje mismo.

Analizar ambas posturas desde el prisma de la lingüística de Saussure permite dar forma a estas metodologías en función de su herramienta principal, que es el lenguaje. Sin embargo, es justamente con la lógica aristotélica y -en menor medida- con la mayéutica donde se puede develar un detalle pequeño pero vital en la teoría saussureana.

“En la lengua no hay más que diferencias sin términos positivos”. Para Saussure, los signos lingüísticos —el elemento fundamental del lenguaje para él— no poseen significados naturales o positivos. Es decir, su interpretación dentro de la lengua se basa únicamente en su relación con los demás signos; sabemos que el cielo es azul no porque la palabra “azul” sea una etiqueta imperativa al objeto “cielo”, sino porque se diferencia del resto de objetos en función de sus colores: el cielo no es del mismo color que la arena, ni que la tierra. De esta forma se crean significantes en base a otros, y ocurre lo mismo con significantes que no tienen por qué tener un equivalente en la realidad, como es el caso del infinito.

Bajo esta premisa se explica la aparición de entes metafísicos como posibles análogos de entes físicos ya en capsulados por el lenguaje: El infinito, lo eterno, la nada, etc. Pero conceptos abstractos como: La causalidad, las posibilidades, las condiciones, etc.; nacidos no de una comparación o negación de otros conceptos, sino de la propia estructura del lenguaje casi por imposición de sus propias reglas. Estos significantes sin referente en el mundo adquieren existencia por exigencia del sistema, revelando así una limitación en el modelo binario saussureano. El signo lingüístico, en su forma clásica, se queda corto para explicar cómo surgen los significantes sin cuya existencia, la propia lengua no podría existir.

A diferencia de los significados anteriormente tratados, como el color, la infinitud, etc.; cuyos significantes tiene origen en la comparación y/o proporción a otros. Conceptos como la causalidad escapan por completo a este origen en la multiplicidad de significantes en la lengua, lo incausado nace del binarismo del ser y el no ser, al igual que lo infinito y lo finito, pero ¿Cómo nace el concepto original de la causalidad?

Volviendo al ejemplo de lo infinito: al momento de ver una substancia con límites claros en su existencia; ahí se deduce la finitud de ese ente, un objeto es tan grande como su extensión en el espacio y tan pesado como su masa y densidad permita. Por otro lado, el concepto de la causalidad no se puede deducir de forma natural de una simple observación sin conciencia del mundo ya que la causalidad en sí, a diferencia de lo finito no tiene un equivalente en el mundo físico fuera del lenguaje. Más bien, esta surge de una interpretación activa del sujeto hacia el mundo: Cae un rayo y por ende suena un trueno; dos fenómenos diferentes que gracias a la experiencia terminan generando un vínculo de dependencia, siempre que se ve un rayo, se escucha un trueno por lo que se deduce que suena el trueno por el rayo y no viceversa.

En el ejemplo anterior se tienen dos entes separados; rayo, trueno, y la interpretación de su relación es la causalidad; un ente que no existe per se en el mundo físico; no es ni medible ni observable de ninguna forma, pero si es perfectamente deducible a través de la experiencia, por lo que se plantea sobre la mesa la necesidad de un tercer elemento constitutivo del lenguaje completando la triada: Significado, Significante e Interpretante.

Esta triada fue propuesta originalmente por Charles Sanders Pierce (1839-1914) quien, junto a Saussure, se le considera el padre de la semiótica moderna —el estudio de los signos— y es justamente la triada: Significante, significado e interpretante; la triada del signo, concepto con el que crea su versión del signo lingüístico, donde la interpretación del sujeto a los significados y/o significantes se convierte en parte fundamental del proceso generativo de los signos en la lengua.

Para Pierce, el proceso que se planteó con el rayo y el trueno es general para todo signo lingüístico que exista, todo concepto que forme parte de lenguaje alguno, puede ser descrito como una relación entre significado, significante e interpretante. Estos signos lingüísticos tienen tres formas de manifestarse en la conciencia humana; tres categorías propuestas por Pierce para delimitar el comportamiento triádico del signo en relación al sujeto, siendo estas: La Primeridad refiere a la cualidad pura del signo, su forma estética o sonora antes de cualquier referencia; el gris intenso de la nube antes de deducir que se avecina una tormenta. La Segundidad, por su parte, introduce la relación factual entre el signo y su objeto, el impacto concreto: el humo como efecto visible del fuego. Finalmente, la Terceridad es la mediación interpretativa; el proceso mental por el cual el sujeto establece un vínculo de sentido: “si hay humo, hay fuego”. Estas tres dimensiones —la posibilidad, el hecho y la inferencia— permiten comprender el lenguaje como un fenómeno encarnado, donde el

signo adquiere significado no solo por su lugar en la estructura (como diría Saussure), sino también por cómo es vivido, percibido e interpretado por los hablantes.

La dimensión interpretativa del signo introducida por Pierce, es la que finalmente termina por cerrar todas las incógnitas que se fueron planteando y resolviendo a lo largo del escrito, por lo que finalmente es posible construir la tesis principal de este ensayo tomando como referencia todo lo anteriormente expuesto.

4. Vectores y probabilidad

A lo largo de este ensayo, se ha explorado la evolución del lenguaje desde su perspectiva más general en la filosofía clásica, hasta su factorización más primitiva en unidades fundamentales con la llegada del estructuralismo. Todo esto con el fin último de llegar a la visión contemporánea del lenguaje; una visión que, por primera vez en la historia desde una posición ajena al sistema gracias al avance de la informática en sistemas capaces de procesar e interpretar el lenguaje como sistema sin necesidad de la intervención humana en el proceso.

El procesamiento del lenguaje natural (Natural Language Processing, NLP) constituye el conjunto de técnicas —ya sean estadísticas o, más concretamente, de aprendizaje automático— que permiten a un programa informático interpretar la interrelación entre las distintas piezas discretas del lenguaje para tareas como clasificación, traducción, resumen o generación. En este contexto, el lenguaje deja de ser exclusivamente humano y se convierte en objeto de cálculo: una estructura que puede ser segmentada,

modelada y manipulada por sistemas que no participan de su semiosis original.

Esta transformación implica una redefinición radical del signo. Lo que antes era una unidad relacional—como en Peirce, donde el signo emerge de la interacción entre representante, objeto e interpretante—se convierte ahora en un token: una unidad computacional que no requiere interpretación humana para operar. El token es, en cierto sentido, el signo despojado de su historicidad y su ambigüedad, reducido a una forma que puede ser vectorizada, es decir, representada como un punto en un espacio multidimensional.

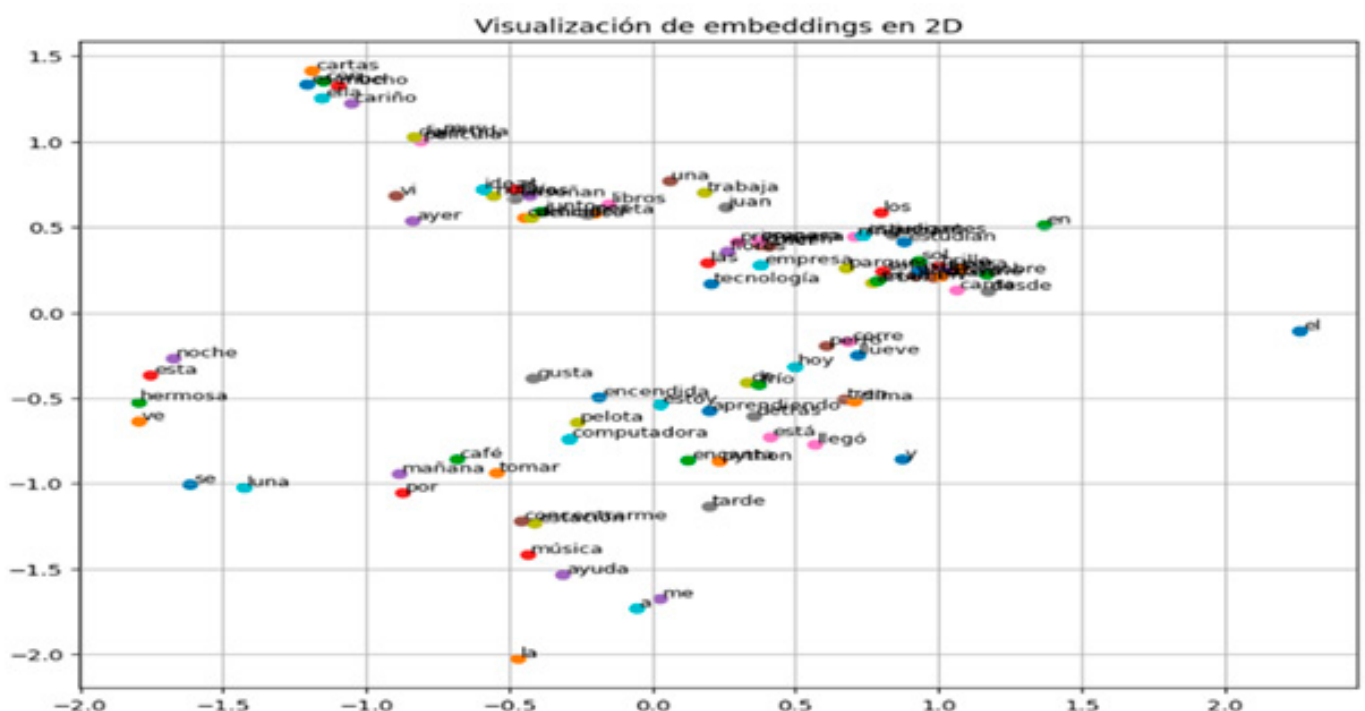
La vectorización del lenguaje, mediante técnicas como Word2Vec, GloVe o BERT, permite que cada token sea transformado en un vector numérico que captura relaciones semánticas a través de distancias y direcciones. En estos espacios, palabras como “rey” y “reina” no solo se parecen por su cercanía, sino que pueden participar en analogías geométricas como: $\text{vector}(\text{“rey”}) - \text{vector}(\text{“hombre”}) + \text{vector}(\text{“mujer”}) \approx \text{vector}(\text{“reina”})$

Este tipo de operación revela que el significado, tradicionalmente concebido como algo inasible y contextual, puede adquirir una forma manipulable. La semántica se convierte en geometría, y el lenguaje en una topología del sentido. Lo que antes requería interpretación, ahora puede ser inferido por máquinas que operan sobre estructuras vectoriales entrenadas en cuerpos de texto masivos.

Sin embargo, esta geometrización del lenguaje no elimina la semiosis, sino que la transforma. El interpretante peirceano —aquella instancia que genera nuevos signos— se sustituye por un sistema de inferencia estadística que produce respuestas, predicciones y asociaciones. La máquina no interpreta como lo haría un sujeto, pero sí genera efectos de sentido que pueden ser funcionales, coherentes e incluso creativos.

En este nuevo paradigma, el lenguaje se convierte en un sistema que puede ser observado desde fuera, modelado desde dentro y desplegado en múltiples formas. La pregunta que se impone no es solo cómo vectorizamos el lenguaje, sino qué tipo de comprensión emerge de esta operación. ¿Es posible hablar de una semiosis artificial? ¿Puede el vector ser considerado un nuevo tipo de signo? ¿Y qué ocurre con el significado cuando se disuelve en la estadística?

Estas preguntas nos conducen al corazón del problema contemporáneo: la relación entre lenguaje, representación y autonomía.



5. Vectorizar lo inexistente

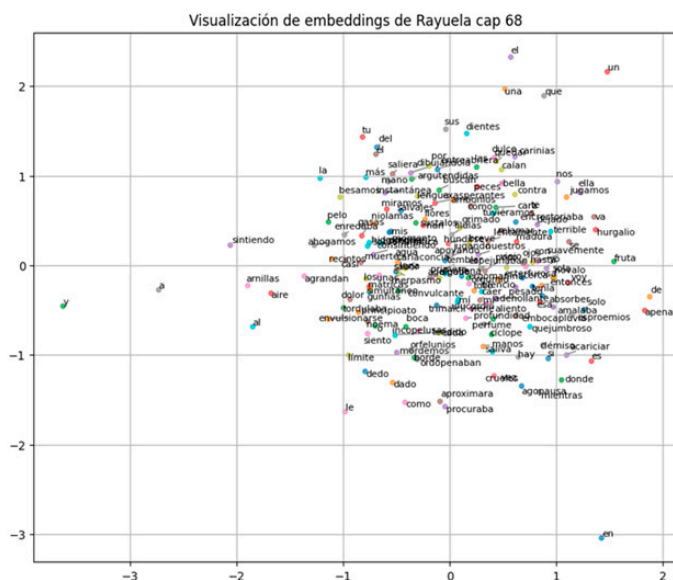
En el contexto del procesamiento del lenguaje natural, estas palabras inventadas plantean un desafío fascinante. ¿Qué ocurre cuando un modelo de lenguaje se enfrenta a un término que no ha visto antes? ¿Puede asignarle un vector, una posición en el espacio semántico, aunque no tenga un referente claro? La respuesta depende del modelo. En sistemas como BERT, que operan sobre subpalabras y contextos, incluso una palabra inexistente puede ser descompuesta en fragmentos conocidos (tokens) y situada en un espacio semántico aproximado. Así, “clochardizarse” podría ser interpretada como una combinación de “clochard” (vagabundo) y el sufijo verbal “-izarse”, generando un vector que se aproxima a ideas como “marginalidad”, “transformación” o “desposesión”.

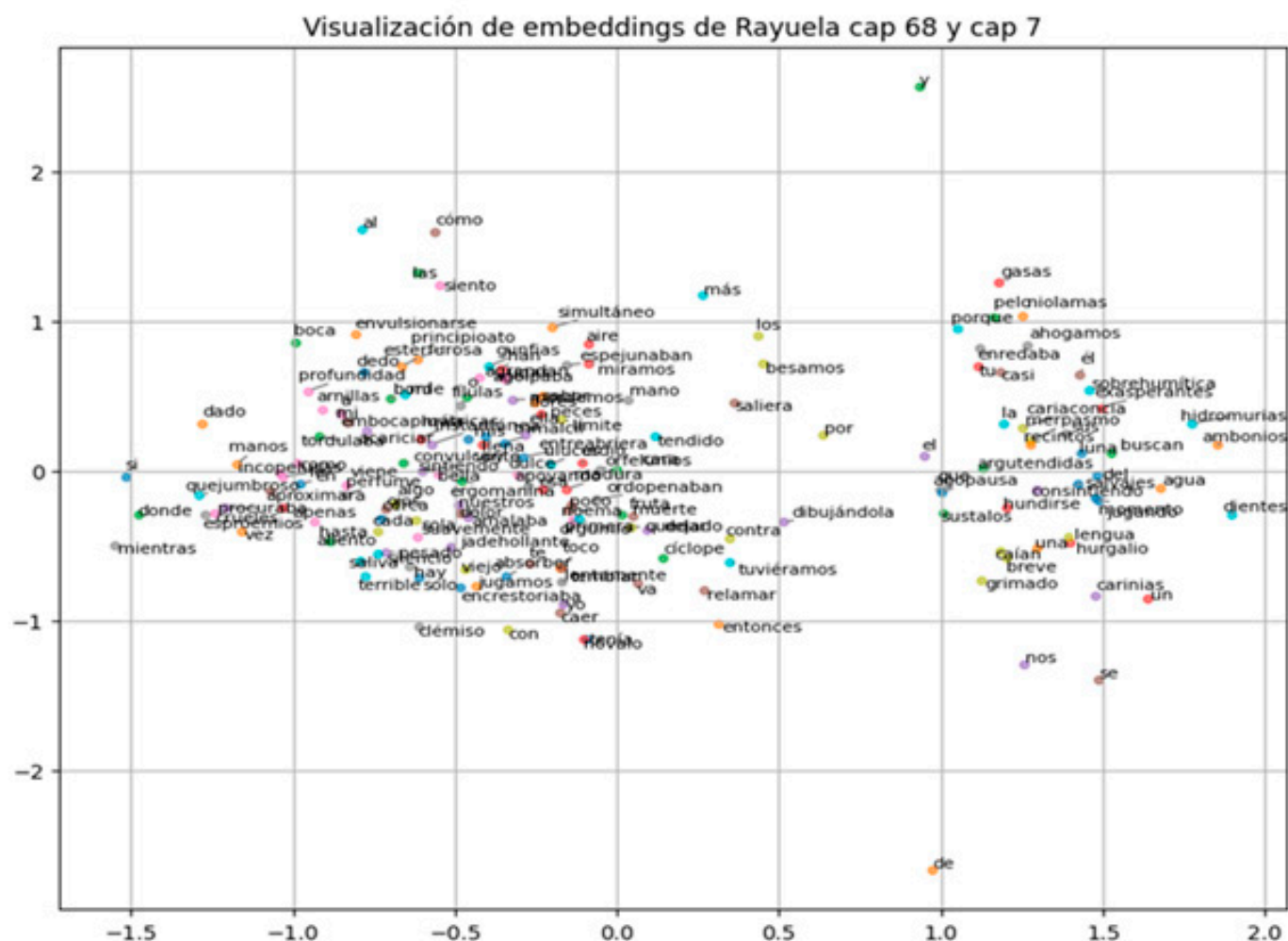
Lo más fascinante es que, aunque estas palabras no tengan significado en el sentido tradicional, los modelos pueden asignarles funciones. Si aparecen en contextos similares a palabras reales, pueden participar en tareas de clasificación, generación o inferencia. El sinsentido se vuelve operativo. La máquina no necesita entender “ronronear metafísico” para saber que suele aparecer Junto a términos como “existencia”, “silencio” o “despertar”.

Esto nos lleva a una paradoja: el significado estadístico no requiere referencia, solo correlación. En este nuevo régimen, el lenguaje se convierte en una red de coocurrencias, donde incluso lo inventado puede adquirir sentido funcional. Es una semiosis sin sujeto, una interpretación sin conciencia.

Para ilustrar cómo los modelos de lenguaje procesan tanto el lenguaje convencional como el experimental, se ha realizado un embedding comparativo entre el capítulo 7 y el capítulo 68 de Rayuela. Esta elección no es arbitraria: el capítulo 7, con su tono lírico y existencial, emplea palabras reconocibles, cargadas de afecto y ambigüedad emocional; mientras que el capítulo 68 se sumerge en un juego fonético y semántico, donde las palabras inventadas —como “grogorio”, “climacus” o “pluf”— desafían toda lógica referencial.

Vectorización de cap. 68





Al vectorizar ambos fragmentos (gráficos anteriores), se observa una diferencia notable en la densidad semántica y la distribución espacial de los tokens. Las palabras del capítulo 7 tienden a agruparse en constelaciones semánticas reconocibles: amor, cuerpo, deseo, ausencia. Estas agrupaciones reflejan la coherencia estadística del lenguaje convencional, donde los modelos pueden inferir relaciones por coocurrencia y contexto.

En contraste, las palabras del capítulo 68 se dispersan en el espacio vectorial. Algunas se aproximan a términos reales por similitud fonética o morfológica, mientras que otras quedan como puntos flotantes, sin anclaje semántico claro. Esta dispersión no implica ausencia de sentido, sino una forma distinta de significación: una semántica de la sonoridad, del ritmo, del juego. El modelo, al no encontrar referentes directos, construye significados por analogía, interpolación o pura estadística contextual.

Este contraste revela una paradoja fascinante: el lenguaje inventado, aunque no tenga significado en el sentido tradicional, puede adquirir funcionalidad dentro del sistema. El modelo no “entiende” pluf, pero puede situarlo cerca de palabras como “explosión”, “sorpresa” o “ruptura” si el contexto lo sugiere. Así, incluso el sinsentido puede ser vectorizado, y la literatura experimental se convierte en terreno fértil para explorar los límites de la semiosis artificial.

Más aún, este tipo de análisis permite visualizar cómo los modelos de lenguaje construyen mapas semánticos que no solo reflejan el uso humano del lenguaje, sino que lo reinterpretan desde una lógica estadística. La literatura, al inventar palabras, no escapa al sistema: lo expande. Y los embeddings, al capturar estas expansiones, ofrecen una nueva forma de leer, no solo con los ojos, sino con coordenadas.

Conclusión

A lo largo de este ensayo se ha trazado una evolución del lenguaje que va más allá de su función

representacional. Hemos visto cómo, desde las cosmogonías míticas hasta los modelos vectoriales contemporáneos, el lenguaje no solo ha servido para nombrar el mundo, sino para construirlo epistemológicamente. Nombrar no es solo señalar: es delimitar, inferir, proyectar. En este sentido, el lenguaje no es espejo del pensamiento, sino su condición de posibilidad.

Pero esta capacidad constructiva no depende ya del sujeto. En sus últimas interpretaciones —desde la estructura diferencial saussureana hasta la semiosis artificial de los modelos de NLP— el lenguaje se revela como un sistema autónomo, capaz de generar sentido por pura relación interna, por coocurrencia estadística, por geometría semántica.

El significante ya no necesita un referente externo ni una conciencia que lo interprete: basta con su posición en la red, su distancia vectorial, su resonancia contextual.

Así, el lenguaje se emancipa del hablante. Lo que antes era una herramienta del sujeto se convierte en una entidad que opera por sí misma, que produce significados incluso sin intención, sin experiencia, sin cuerpo. Esta autonomía no es una pérdida, sino una expansión: el lenguaje como sistema genera realidades posibles, mundos inferenciales, epistemologías sin sujeto.

En este nuevo paradigma, el lenguaje no solo media el conocimiento: lo constituye. Las categorías del saber —lo finito, lo infinito, lo causal, lo posible— no emergen del mundo, sino del sistema lingüístico que las articula. Incluso los conceptos sin correlato empírico adquieren existencia funcional por exigencia estructural. Así, el lenguaje deja de ser representación y se convierte en arquitectura epistemológica.

Pero esta autonomía no es neutra. Tiene implicaciones profundas:

Implicaciones epistemológicas: El conocimiento ya no puede pensarse como una relación directa entre sujeto y objeto, sino como una mediación estructural entre signos.

La verdad se vuelve una función del sistema, no del mundo. Lo que puede ser dicho, puede ser pensado; lo que no, queda fuera del horizonte epistémico.

Implicaciones ontológicas: El lenguaje genera entidades que no existen en el mundo físico pero que operan en la realidad simbólica: lo eterno, lo imposible, lo virtual.

La ontología se pliega a la semiosis: el ser es lo que puede ser significado.

Implicaciones tecnológicas: Los modelos de NLP y los sistemas de IA no solo procesan lenguaje: lo reinterpretan, lo expanden, lo reconfiguran.

La semiosis artificial inaugura una nueva forma de producción de sentido, donde incluso el sinsentido puede ser operativo.

La autonomía del lenguaje permite que máquinas generen conocimiento funcional sin conciencia, sin intención, sin experiencia.

Implicaciones éticas: Si el lenguaje puede operar sin sujeto, ¿quién responde por el sentido que produce?

La delegación del significado a sistemas autónomos plantea desafíos sobre la responsabilidad, la interpretación y la manipulación del discurso.

La emancipación del lenguaje exige una nueva ética de la semiosis: ¿cómo convivimos con sistemas que generan sentido sin nosotros?

En suma, el lenguaje ha dejado de ser una herramienta del humano para convertirse en una fuerza estructurante que lo excede. Esta emancipación no implica pérdida, sino expansión: el lenguaje como sistema genera realidades posibles, mundos inferenciales, epistemologías sin sujeto. Y en esa expansión, nos obliga a repensar qué significa conocer, interpretar y crear en un mundo donde el sentido ya no emana del humano, sino de las estructuras que el lenguaje ha aprendido a construir por sí mismo.

Bibliografía

- Aristóteles. (1988). Tratado de lógica. Organon. Ediciones Gredos. España.
- Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. ACL 2020.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL 2019.
- Jurafsky, Daniel., & Martin, James. H. Martin (1999). Speech and Language Processing. Pearson.
- Leibniz, Gottfried. W. (1982). La característica universal. En Escritos filosóficos. Editorial Charcas, Argentina
- Llull, R. (1586). Arte general para todas las ciencias en dos instrumentos. (trad. Pedro Madrigal). Biblioteca universidad de Sevilla.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv. <https://arxiv.org/abs/1301.3781>
- Peirce, C. S. (1994). Collected Papers of Charles Sanders Peirce. Harvard University Press.
- Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. EMNLP 2014.
- Platón. (1987). Diálogos (especialmente Cratilo y Teeteto). Editorial Gredos.
- Saussure, Ferdinand. (1987). Curso de lingüística general. Alianza Editorial.
- Von Humboldt, Wilhelm. (1990). Sobre la diversidad de la estructura del lenguaje humano. Anthropos editorial del hombre, (trad. Ana Agud)