

Este artículo revisa los modelos masivos de lenguaje que han aparecido en los últimos tiempos desde el punto de vista de una lingüista. De forma divulgativa, en este trabajo se pretende mostrar que, si bien no puede ponerse en duda la utilidad de estas nuevas herramientas, no es cierto que sean generadoras de lenguaje, al menos de la manera en que lo hace un ser humano.

La aparición de modelos masivos del lenguaje en los últimos tiempos ha supuesto una revolución en diversos ámbitos, entre los que cabe destacar la Lingüística. Si las máquinas son capaces de hablar, puede pensarse que estas son cada vez más humanas y, en caso de que esto sea cierto, es posible que en el futuro sean capaces de pensar por sí mismas y tomar decisiones.

No pretendo en este artículo aclarar si lo último es cierto, pero sí negar que hoy por hoy estos modelos sean generadores de lenguaje natural. Así, en este artículo mostraré, de manera necesariamente breve e incompleta, qué es el lenguaje humano y qué son las lenguas para establecer después una escueta comparación con estos modelos cuya implantación más conocida es ChatGPT.

La arquitectura del lenguaje humano

He de advertir antes de empezar que el punto de partida es, como el de todo investigador, sesgado. El acercamiento científico a cualquier objeto de estudio del mundo natural, como es el lenguaje (y, por extensión, a las lenguas) ha de hacerse asumiendo una teoría compartida por un conjunto amplio de científicos. En mi caso, asumo las propuestas generativistas, cuyo mayor exponente es el lingüista y pensador Noam Chomsky, que entienden que los humanos estamos dotados de una especie de dispositivo que nos predispone genéticamente para adquirir una o más lenguas desde el momento de nuestro nacimiento; este mecanismo se conoce como Facultad del Lenguaje. Siendo esta capacidad inherente al ser humano y similar en todos nosotros, cabe preguntarse por qué hablamos distintas lenguas. La propuesta generativista es que las diferencias entre las más de 7000 lenguas que se hablan en nuestro planeta son solamente superficiales, siendo mucho más aquello que comparten que lo que las diferencia. A modo de ejemplo, todas las lenguas tienen oraciones y todas las oraciones han de tener sujeto; lo que diferencia a este respecto a unas lenguas de otras es que en unos casos el sujeto ha de ser un sintagma explícito y en otros puede no pronunciarse bajo ciertas circunstancias.

La aparición de estos modelos supone un avance enorme para el acceso al conocimiento

Este planteamiento biológico del lenguaje humano se completa con un componente social: para desarrollar una lengua es necesario entrar en contacto con ella en los primeros años de nuestra vida. Por eso hablamos de lenguas maternas. En teoría, no hay un límite en el número de lenguas maternas que un individuo puede adquirir (por eso la mayoría de los hablantes de nuestro planeta son multilingües); la condición *sine qua non* para hacerlo es estar en contacto suficiente con esa lengua, para que nuestra mente pueda “descubrir inconscientemente” (dicho de forma muy simplificada) cuáles son los parámetros de esta (es decir, para que la mente del individuo pueda determinar el sistema) y así poder desarrollarla.

Esta teoría del lenguaje, que es también una teoría de la mente, no es la única manera que tenemos los lingüistas de acercarnos a nuestro objeto de estudio. En el lado opuesto, digamos, se sitúan aquellos científicos que proponen que el lenguaje ha de ser estudiado a través de sus manifestaciones, es decir, las producciones lingüísticas.

La razón principal para optar por los supuestos generativistas es que asumir que estamos genéticamente dotados de un órgano del lenguaje nos permite explicar por qué un niño de tres años es capaz de emitir oraciones largas y de una complejidad considerable cuando no puede siquiera comprender el mecanismo de la suma. Puede objetarse que el lenguaje forma parte de la inteligencia general y que esta es la razón por la que desarrollamos una lengua (porque vamos desarrollando nuestra inteligencia), pero entonces tendríamos que explicar por qué hay individuos que sufren trastornos del lenguaje, como el trastorno específico del lenguaje (TEL) o los distintos tipos de afasias, mientras mantienen intacta su inteligencia general, y también por qué hay individuos que tienen seriamente dañada la inteligencia general mientras que presentan un desarrollo normal del lenguaje, como aquellos que padecen Síndrome de Williams.

Partimos, pues, de que el lenguaje es una capacidad de la especie humana para la que estamos dotados genéticamente. Asumimos también que, durante sus primeros años de vida, el individuo debe estar expuesto a una o varias lenguas naturales que le permiten desarrollar una o varias lenguas. Por último, digamos que cuando este periodo de adquisición de la lengua materna finaliza (allá por la pubertad) el individuo es competente en dicha lengua (y, por tanto, todos los individuos somos “expertos” en nuestra lengua materna).

El componente sintáctico como elemento central de una teoría generativa

El siguiente paso consiste en explicar cómo producimos oraciones una vez que somos competentes. El modelo de T invertida, que es el que manejamos los generativistas, propone lo siguiente (de manera muy simplificada): el ser humano adquiere a través del contacto directo con una lengua las palabras de esta (siendo rigurosa, no debería hablar de palabras sino de ítems léxicos, que no son conceptos siempre equivalentes). Estas palabras contienen toda la información lingüística relevante para que nuestra mente sepa cómo utilizarlas. Por ejemplo, una palabra como *comer* tendrá, al menos, la siguiente información: es un verbo, se pronuncia /kómer/, su significado aproximado es “ingerir alimentos” y, cuando se inserta en la sintaxis, debe tener un sujeto (que es el agente) y un objeto (que es el tema).

Supongamos que nuestro cerebro va a formar la oración *Los niños comieron pasta con tomate*. Lo primero que hacemos es seleccionar de nuestro léxico un subconjunto de palabras, en este caso “los”, “niños”, “comer” [flexionado en pasado, tercera persona y plural], “pasta”, “con” y “tomate”. A continuación, insertamos estos elementos en el componente sintáctico, que es central en esta teoría. En un modelo de este tipo el ensamble de piezas es binario, de modo que se producirá, aproximadamente, de la siguiente manera: “con” se une a “tomate” (y forman un sintagma preposicional), el resultado se une a “pasta” (sintagma nominal); el nuevo sintagma se une al verbo del que es complemento (“comen pasta con tomate”); por otro lado, “los” y “niños” se han unido formando un sintagma nominal. Por último, el sintagma “los niños” se une al sintagma verbal “comieron pasta con tomate” y tenemos una oración bien formada en el componente sintáctico.

A partir de aquí la sintaxis envía el resultado a dos interfaces: la fonético-fonológica, encargada de producir una secuencia sonora (o gestual en el caso de las lenguas de signos), y la interfaz de forma lógica, encargada de interpretar semánticamente el objeto lingüístico creado.

Cualquier “complicación” a lo largo de este proceso (por ejemplo, que aparecieran dos verbos) produciría que “la derivación se estrellara”; es decir, que se generaría una secuencia agramatical que no percibimos simplemente porque no llega a ser generada.

Estas IAs no son capaces de entender ni lo que dicen ni lo que se les dice

La explicación que acabo de dar está enormemente simplificada y es incompleta. Me interesa, no obstante, destacar para lo que sigue que esta concepción del lenguaje no es posible si no se asume (como de hecho se hace en todas las teorías actuales sobre el lenguaje humano) que la sintaxis es esencialmente jerárquica y no lineal (es decir, que las palabras se van agrupando de dos en dos en una suerte de árbol invertido).

La generación del lenguaje por parte de máquinas

Una vez que he formulado los principios básicos de cómo se articula el lenguaje humano, pensemos en qué hacen y cómo lo hacen los modelos de lenguaje natural. Adelanto que no soy una experta en el tema, pero lo que parece evidente es que estos modelos son capaces de elaborar textos con una cohesión y coherencia tan apabullantes que, realmente, pareciera que hablaran. Lo que quizás no resulte tan evidente es que no tienen ni la más remota idea de lo que están diciendo, por un lado, y que el hecho de que sus producciones sean gramaticales se debe a una cuestión matemática y no lingüística. Me explico a continuación.

He mencionado anteriormente que el punto de partida en la formación de todas nuestras producciones lingüísticas son los ítem léxicos o palabras. ChatGPT no trabaja con palabras ni con ítems léxicos, sino con lo que se denomina *tokens* que son, *grosso modo*, conjuntos de letras (dado que es un modelo escrito) que se combinan matemáticamente. Según cuenta de manera magistral José Luis Mendivil en el artículo «Un loro escolástico en la habitación china: ¿qué nos enseña ChatGPT sobre la mente humana»¹, estos modelos de generación de lenguaje se limitan a crear, de forma lineal (insisto en que crean de forma lineal y no jerárquica, como lo hace el lenguaje humano), cadenas del lenguaje a partir de modelos matemáticos; es decir, que analizan un conjunto enorme de textos para ser capaces de mostrar qué palabra aparecerá detrás de la que ya han escrito.

Sin duda, la aparición de estos modelos supone un avance enorme para el acceso al conocimiento. ChatGPT es capaz de explicarnos de forma muy detallada conceptos sencillos y cuya información está recogida en algún punto de los conjuntos de textos con los que trabaja. Lo que ya no hace tan bien este modelo de lenguaje (ni tan bien ni tan mal porque directamente no lo hace) es crear nuevas emisiones o interpretar las producidas y, mucho menos, ser capaz de determinar si una producción es o no posible en una lengua. Pondré un ejemplo. Le he preguntado a ChatGPT lo siguiente (con la respuesta que hay en la misma foto):



¿Puedes decirme cuál es la diferencia entre "un joven viejo" y "un viejo joven"? ¿O son sinónimas?



"Un joven viejo" y "un viejo joven" son expresiones que se utilizan para describir a una persona de una manera un tanto peculiar. Aunque no son sinónimas, ambas sugieren una contradicción en la edad aparente y el comportamiento de la persona:

- "Un joven viejo" se refiere a alguien que, a pesar de tener una edad avanzada, actúa o se comporta de una manera más joven de lo que cabría esperar. Es decir, es un anciano que mantiene una actitud juvenil o intereses propios de personas más jóvenes.
- "Un viejo joven" se refiere a alguien que, a pesar de ser joven en términos de edad cronológica, se comporta o tiene actitudes típicas de personas mayores. Es decir, es un joven que actúa como si fuera más viejo de lo que es.

En resumen, la diferencia está en la edad cronológica y el comportamiento. "Un joven viejo" es mayor en edad pero actúa joven, mientras que "un viejo joven" es joven en edad pero actúa viejo.

Pregunta a ChatGPT. Fuente: Silvia Gumiel.

Cómo se puede observar por la respuesta, este modelo del lenguaje no es capaz de comprender la diferencia entre dos emisiones que un niño de seis años entiende perfectamente (su respuesta es errónea, la interpretación que hace un humano es exactamente la contraria). Y la razón es fundamentalmente que las IAs de procesamiento de textos no entienden ni lo que están diciendo ni cómo están construyendo su "lenguaje" mientras que un ser humano, de manera inconsciente, sabe que un joven viejo, es un sintagma nominal cuyo núcleo es "joven" (por tanto, un joven viejo es un tipo de "joven") mientras que en un viejo joven el núcleo es "viejo" y, por tanto, nos referimos a un tipo de "viejo".

Avancemos un poco más. Le he pedido a ChatGPT que cree una oración agramatical. Esta ha sido su respuesta:



¿Puedes crear una oración agramatical?



Sí, puedo crear una oración agramatical. Aquí tienes un ejemplo:

"El pájaros volar."

Esta oración no es gramaticalmente correcta porque el verbo "volar" debería concordar en número con el sustantivo "pájaros," y la forma correcta sería "Los pájaros vuelan."

Pregunta a ChatGPT. Fuente: Silvia Gumiel

Aunque *a priori* podría parecer que tiene razón, lo cierto es que «los pájaros volar» podría ser una oración bien formada en algunos contextos, aunque no haya concordancia entre el sujeto y el verbo (que no puede concordar porque está en infinitivo) como, por ejemplo: me gusta ver [los pájaros volar]. Lo que, de hecho, hace agramatical esa secuencia es que el determinante no concuerda con el nombre al que modifica.

Sirvan estos dos ejemplos para mostrar que estas IAs no solo no hablan, sino que no son capaces de entender ni lo que dicen ni lo que se les dice. Con toda seguridad, estos modelos se irán mejorando, pero creo que puedo afirmar sin equivocarme que, mientras se basen en modelos matemáticos y su implementación no esté fundamentada en los principios que articulan la arquitectura del lenguaje humano, no será posible que genere lenguaje natural y, mucho menos, que lo entienda.

Esto no quiere decir que estos modelos no sean un instrumento fantástico para avanzar en el conocimiento. También lo serán para retroceder si no los utilizamos adecuadamente. Me remito, por ejemplo, a la reciente noticia de una investigadora que presume de haber ganado una beca de creación de la Generalitat para una investigación artística creando su memoria con ChatGPT. No conozco los detalles, pero me atrevo a afirmar que, o bien la autora tomó el texto generado en ChatGPT como base y, a partir de ahí, añadió lo fundamental (la aportación humana que haya hecho a ese proyecto merecedor de la beca), o bien que los evaluadores no hicieron bien su trabajo. Y me inclino más por lo primero que por lo segundo.