

Análisis crítico de la Inteligencia Artificial Generativa en teorías del lenguaje

Critical analysis of Generative Artificial Intelligence in theories of language

Yolima Janeth Rodríguez Jiménez¹ (yoli.rodriquez@udla.edu.co) (<https://orcid.org/0000-0003-3108-1600>)

Frank Albert Pama Arias² (f.pama@udla.edu.co) (<https://orcid.org/0009-0006-0547-3140>)

Resumen

La inteligencia artificial generativa (GAI) ha redefinido la producción de lenguaje, influyendo en ámbitos que van desde la educación hasta la creación de contenido. Sin embargo, su evolución no es únicamente un logro computacional; está profundamente arraigada en teorías lingüísticas que moldean sus capacidades y limitaciones. Este artículo examina críticamente la relación simbiótica pero asimétrica entre la lingüística y la IAG, destacando cómo los principios lingüísticos informan el diseño de modelos mientras son instrumentalizados sin reflexión ética o cognitiva. Desde un enfoque de investigación cualitativa con un alcance crítico-analítico, este estudio realiza una revisión sistemática de la literatura reciente, en donde compara perspectivas teóricas y empíricas sobre la dependencia de los marcos lingüísticos por parte de la IAG. A través del análisis del discurso, explora dos dimensiones clave: (1) la apropiación selectiva de teorías lingüísticas en la arquitectura de modelos y (2) las preocupaciones éticas en la generación de contenido. Los hallazgos indican que, aunque modelos como GPT-4 y BERT integran sintaxis y semántica, no logran captar las sutilezas pragmáticas y socioculturales, lo que refuerza las inequidades sistémicas. El estudio concluye que la futura investigación interdisciplinaria ha de priorizar marcos éticos de IA, la diversidad lingüística y la modelización cognitiva para desarrollar sistemas generativos más inclusivos y responsables. Este enfoque crítico subraya la necesidad de ir más allá de la mera optimización tecnológica hacia un paradigma de IA más equitativo y consciente lingüísticamente.

Palabras clave: lingüística computacional, inteligencia artificial generativa, sesgo lingüístico, ética del PLN, modelado cognitivo.

Abstract

Generative artificial intelligence (GAI) has redefined language production, influencing fields ranging from education to content creation. However, its evolution is not solely a computational achievement; it is deeply rooted in linguistic theories that shape its capabilities and limitations. This article critically examines the symbiotic but

¹ Magíster en Educación Universidad Libre y Universidad de Las Tunas. Licenciada en lingüística y literatura. Profesora asociada y Líder semillero Texto y voz. Universidad de la Amazonia. Colombia.

² Ingeniero en Informática y Sistemas. Analista de Datos Google. Science Date IBM. Diseñador y Productor Multimedia. Esp. Seguridad Informática. Esp. en IAG. Esp. Marketing digital. Semillerista de Texto y voz de la Licenciatura en Literatura y Lengua Castellana. Universidad de la Amazonia. Colombia.

asymmetrical relationship between linguistics and AGI, highlighting how linguistic principles inform model design while being instrumentalized without ethical or cognitive reflection. From a qualitative research approach with a critical-analytic scope, this study conducts a systematic review of recent literature, where it compares theoretical and empirical perspectives on IAG's reliance on linguistic frameworks. Through discourse analysis, it explores two key dimensions: (1) the selective appropriation of linguistic theories in model architecture and (2) ethical concerns in content generation. Findings indicate that although models such as GPT-4 and BERT integrate syntax and semantics, they fail to capture pragmatic and sociocultural subtleties, reinforcing systemic inequities. This critical approach underscores the need to move beyond mere technological optimization towards a more equitable and linguistically aware AI paradigm.

Key words: computational linguistics, generative artificial intelligence, linguistic bias, PLN ethics, cognitive modeling.

Introducción

Desde los aportes de Chomsky sobre el aspecto creador del lenguaje, que ya habían sido motivo de interés en Descartes, Cordemoy, Harris, Schlegel, Humboldt y la escuela de Port Royal, se revisó cómo desde medios finitos se generan mensajes infinitos (Bernal Leongómez, 1984). De esta manera es como desde el componente sintáctico se divide en dos subcomponentes, el primero categorial compuesto por reglas de ramificación y subcategorización, y el segundo el lexicón, que corresponde al conjunto de artículos léxicos que configuran el vocabulario de una lengua, y estos dos se especificarán con características fonológicas, sintácticas y semánticas.

Es desde los cambios de la estructura profunda a la superficial que se sitúa uno de los aspectos que soportan los avances de la IAG. Se trata pues de las reglas de transformación, en donde es posible hacer permutación, adición, elisión y sustitución. Así, modelos como GPT-4 o BERT no solo procesan texto, sino que encapsulan teorías sobre sintaxis, semántica y pragmática desarrolladas décadas atrás (Devlin et al., 2019). No obstante, esta interdependencia entre lingüística e IAG rara vez se examina de forma crítica.

Sin embargo, esta relación simbiótica es asimétrica en tanto la lingüística aporta las reglas y la IAG las instrumentaliza sin reflexionar sobre sus implicaciones éticas o cognitivas. Si bien desde la gramática generativa transformacional y otros estudios desde la lingüística se plantearon los universales lingüísticos, los cuales sugieren que todas las lenguas son similares y sus variaciones radican en la estructura superficial³, se mantuvo el interés de llevar esto a la automatización. Sabemos el diseño de la gramática de estado finito fue un reto del pasado que Chomsky hizo a partir del modelo

³ Greemberg (1963) ofrece cuarenta y cinco universales, quien se basa en las nociones de universal relativo y de universal implicativo para ofrecer una nueva perspectiva en la búsqueda de universales.

de Shannon y Weaver de la teoría de la comunicación que permitía la generación inmediata y automática de un mensaje desde un proceso estocástico, que en matemáticas se denomina proceso de Markov que consideraba la comunicación como un proceso lineal. Con la creación de las IA, esto ha derivado en sistemas que, aunque eficientes, perpetúan sesgos culturales, desatienden lenguas minoritarias y trivializan el plagio académico (Gallent-Torres et al., 2023; Névéol et al., 2018a).

Ante este panorama, este artículo cuestiona ¿cómo ha moldeado la lingüística el desarrollo técnico de la IAG? ¿De qué modo sus omisiones teóricas en pragmática y variación dialectal limitan su potencial? A través de un análisis crítico, se exploran dos dimensiones: (1) la apropiación selectiva de teorías lingüísticas en el diseño de modelos y (2) los desafíos éticos en la generación de contenido. Solo al deconstruir estos ejes se podrá avanzar hacia una IAG ética, inclusiva y cognitivamente robusta.

Materiales y métodos

El estudio adoptó un enfoque cualitativo con un alcance crítico-analítico, y se centró en explorar cómo los principios lingüísticos subyacen al diseño y funcionamiento de los modelos de inteligencia artificial generativa (IAG). Este enfoque permitió examinar no solo las dimensiones técnicas de la IAG, sino también sus implicaciones éticas, sociales y culturales.

Se recolectaron más de 200 artículos preliminares, los cuales se filtraron por título y resumen, y se redujeron a 75 estudios que se revisaron para verificar su relevancia y calidad metodológica, lo que dejó 50 artículos clave. Los documentos analizados incluyen artículos académicos, libros, capítulos de libros y documentos de conferencias publicados entre 2018 y 2024. Las fuentes fueron seleccionadas de bases de datos como Web of Science, Scopus, PubMed, Google Scholar y repositorios especializados.

Resultados y discusión

El análisis crítico del discurso se realizó en varias etapas, en primer lugar, se realizó una lectura exhaustiva de los documentos seleccionados para identificar temas, conceptos clave y patrones discursivos. De esta forma, se aplicó un enfoque inductivo para categorizar los datos en temas emergentes, como “apropiación selectiva de teorías lingüísticas” y, “sesgos culturales”. Luego, se compararon los hallazgos entre los diferentes documentos para identificar convergencias y divergencias en las perspectivas presentadas y finalmente se integraron los resultados.

Para garantizar la rigurosidad del análisis, se utilizaron herramientas y técnicas específicas, así, para la gestión bibliográfica se empleó Mendeley y para el análisis se usó AtlasTi. Como limitantes se tuvieron en cuenta solo documentos en inglés y español, y en algunos casos, los estudios no tenían acceso abierto.

El análisis crítico del discurso se basó en los principios propuestos por Fairclough (1992) y Van Dijk (1993), haciendo énfasis en dos aspectos, en primer lugar, las narrativas dominantes sobre IAG con la identificación de patrones discursivos

recurrentes, como la instrumentalización de teorías lingüísticas y, en segundo lugar, se analizó cómo los autores articulan sus argumentos, al usar marcos teóricos como la lingüística crítica y la ética en IA. El análisis crítico del discurso permitió una revisión exhaustiva y reflexiva de la literatura, de forma que garantice que los hallazgos reflejen tanto los avances técnicos como las preocupaciones éticas y sociales asociadas con la IAG.

Inteligencia Artificial Generativa (IAG) y lingüística

La relación entre la lingüística teórica y la inteligencia artificial generativa (IAG) es un diálogo asimétrico pero indisociable. Desde los postulados de Noam Chomsky en los años 50, la gramática generativa ha buscado describir las estructuras innatas que subyacen al lenguaje humano, proponiendo reglas universales para generar oraciones gramaticalmente válidas (Chomsky, 1965). Este enfoque, centrado en la sintaxis y la recursividad, no solo revolucionó la lingüística, sino que sentó las bases conceptuales para los modelos de lenguaje actuales.

La investigación de Newell y Simon para modelar el razonamiento humano mediante computadora se inspiró en las ideas y el argumento de Chomsky sobre el lenguaje formal. Chomsky coloca las primeras semillas de los lenguajes de programación, y Newell & Simon las primeras semillas de la IA (Mohammed y Shormani, 2024). Como se ve, Chomsky con la lingüística generativa supone que el lenguaje es un fenómeno gobernado por reglas, y que el cerebro humano puede generar un conjunto infinito de oraciones siguiendo dichas reglas.

En este sentido, un lenguaje no verbal, como los lenguajes de programación Python depende de la sintaxis y la semántica debido al hecho de que, si se viola algún principio, se genera una notificación de un error de sintaxis o semántico (Jackson, 2018).

Sin embargo, la transición de la teoría a la práctica en la IAG ha estado marcada por simplificaciones, omisiones críticas y una instrumentalización selectiva de los principios lingüísticos. Por ejemplo, los transformers, arquitectura detrás de modelos como BERT y GPT, replican de forma implícita la idea de “competencia lingüística” propuesta por Chomsky. Al procesar texto mediante atención bidireccional, estos modelos capturan dependencias jerárquicas entre palabras, lo que simula la capacidad humana para reconocer estructuras sintácticas complejas (Devlin et al., 2019). Por ejemplo, en frases como “El libro que el estudiante que aprobó el examen escribió fue publicado”, los transformers identifican relaciones anidadas (incrustación de cláusulas) sin requerir reglas explícitas, algo que la gramática generativa atribuye a un dispositivo cognitivo innato.

No obstante, esta similitud es superficial, pues mientras Chomsky postulaba una “gramática universal” biológicamente determinada, los transformers aprenden patrones estadísticos a partir de datos masivos, sin internalizar reglas abstractas (Greco & Tagarelli, 2024). Es decir, su “competencia” es una ilusión emergente de la exposición a

billones de tokens, no una representación de mecanismos cognitivos. Esto plantea una paradoja: aunque la IAG logra generar lenguaje fluido, su comprensión es frágil y contextualmente limitada. Y es que, aunque la IA busca disminuir la distancia mediante el desarrollo de modelos de datos, algoritmos de aprendizaje profundo y redes neuronales que imitan la inteligencia y el lenguaje humanos (Pannachin y Goertzel, 2007), falta mucho para autenticar y aceptar la IA como una representación de la inteligencia humana.

Ante este dilema, se encuentra la postura de que “la división entre inteligencia natural e inteligencia artificial desaparecerá en la inteligencia híbrida aumentada” (Pan, 2016, citado por Beerends & Aydin, 2024) pero se desconocen las implicaciones éticas y la forma como la IAG presenta falencias. Por ejemplo, un estudio de Riccardi et al. (2024) con el Two Word Test reveló que modelos como GPT-4 fallan al discernir combinaciones léxicas absurdas (p. ej., "cabra cielo"), lo que deja ver que su "comprensión" se reduce a correlaciones estadísticas, no a una representación semántica sólida. Al respecto, Watson (2019) sugiere que el impulso por humanizar los algoritmos es un obstáculo para conceptualizar adecuadamente los desafíos éticos que plantean las tecnologías emergentes y cuestiona las comparaciones o afirmaciones antropomórficas del algoritmo de red neuronal, cuyas similitudes con la cognición humana, están enormemente exageradas.

Y, es que la gramática generativa fue un punto de partida crucial para la IAG, pero su legado se ha diluido en un mar de datos y optimizaciones estadísticas. Para que la próxima generación de modelos trascienda la mera imitación superficial, es imperativo hacer un reenfoque interdisciplinario: los lingüistas deben participar activamente en el diseño de arquitecturas, y los ingenieros deben abandonar la ilusión de que el lenguaje puede reducirse a vectores. Solo así la IAG podrá aspirar a una comprensión genuina, no estadística, del lenguaje humano.

Ahora bien, en un estudio de Li, Nye y Andreas (2021) encontraron que la predicción en modelos de lenguaje neuronal previamente entrenados está respaldada por representaciones dinámicas de significado y una simulación implícita del estado de entidad, y que este comportamiento se puede aprender solo con texto como datos de entrenamiento, es decir, tienen similitudes funcionales con los modelos lingüísticos de semántica dinámica.

Así, el aprendizaje por transferencia, en el que un modelo primero se entrena previamente en una tarea rica en datos antes de ajustarse en una tarea posterior, se ha convertido en una técnica poderosa en el procesamiento del lenguaje natural (PLN). El trabajo en equipo de Colin (2020) logró abordar el problema de texto de entrada y texto nuevo como salida. Sin embargo, hay una desconexión epistemológica entre la IAG y la fundamentación lingüística respecto a cómo los humanos entienden y usan el lenguaje.

Se trata pues de pensar y retomar los estudios acerca del lenguaje y pensamiento. De acuerdo con la formulación más reciente de la gramática generativa, biolingüística, y se piensa desde la pregunta ¿cuál es su diseño básico? de cómo evoluciona el lenguaje y

cuáles son los correlatos neuronales del lenguaje, donde se deduce que el lenguaje humano se basa en un mecanismo computacional particular, realizado neuronalmente, que produce una gama infinita de expresiones estructuradas (Berwick & Chomsky, 2013).

Así es como, en la búsqueda de ofrecer una mejora en la IAG, se encuentra SemBERT, que integra etiquetado de roles semánticos (SRL) en BERT, con el fin de mejorar la comprensión de predicados y argumentos (Zhang et al., 2023) y, los modelos ecológicos, que buscan incorporar la variación dialectal y el contexto cultural con el fin de evitar sesgos hegemónicos (Inoue, 2024). Sin embargo, se ha presentado la generación de contenido sesgado o fraudulento.

Además, la expansión de los modelos de lenguaje, como GPT-3, ha revelado una notable mejora en su rendimiento, especialmente en tareas que requieren independencia y en situaciones donde hay pocos ejemplos disponibles para el aprendizaje (Brown et al, 2020), pues se basa en una arquitectura de transformadores y cuenta con la impresionante cifra de 175 mil millones de parámetros, lo que lo convierte en un modelo diez veces más grande que cualquier otro anterior que no empleara dispersión. Esta considerable cantidad de parámetros permite al modelo manejar y analizar grandes volúmenes de datos, identificando patrones complejos dentro del contexto. Por lo tanto, GPT-3 es capaz de generar texto coherente y relevante, sin necesidad de ajustes extensivos.

Cuestiones éticas y del lenguaje

Frente al incremento en el uso de la inteligencia artificial generativa (IAG), impulsado por el auge de los beneficios que ofrece, se evidencia cómo estas tecnologías no solo replican patrones lingüísticos complejos, sino también sesgos sociales profundamente arraigados. El uso generalizado y, en muchos casos, desinformado de herramientas como ChatGPT o Gemini revela una tendencia creciente hacia la automatización de procesos cognitivos.

Los usuarios buscan principalmente el ahorro de tiempo y la elaboración rápida de tareas, recurriendo a la réplica mecánica de fórmulas de prompts, sin reflexionar sobre las implicaciones éticas, sociales y cognitivas que esto conlleva. Esta práctica contribuye a una aproximación superficial al conocimiento, donde el pensamiento crítico se ve reemplazado por la simple repetición de instrucciones, generando así un uso acrítico de tecnologías que moldean el discurso y la información.

Según Kessler y Casal (2024) y Névéol et al. (2018b), los modelos generativos como ChatGPT y Gemini reflejan desigualdades históricas, ya que se entrenan en datos que priorizan lenguas hegemónicas y perspectivas dominantes, lo que perpetúa la marginalización de comunidades lingüísticas minoritarias. Esta reproducción de desigualdades se manifiesta también en estereotipos de género, raza y clase, lo que plantea una serie de retos éticos y sociales urgentes.

En este sentido, el fenómeno del uso desinformado de IAG afecta directamente los procesos cognitivos, fomentando una dependencia tecnológica que impide el desarrollo de habilidades críticas y analíticas. La facilidad con la que se accede a estas herramientas y la velocidad con la que generan contenido de aparente calidad fomentan una confianza ciega en los resultados, lo que limita la capacidad de los usuarios para evaluar, interpretar y contextualizar la información. Este escenario promueve la replicación de narrativas dominantes sin cuestionamiento alguno, reforzando así estructuras de poder que favorecen a quienes controlan la producción y distribución del conocimiento. En este sentido, la IAG se convierte en un mecanismo que no solo automatiza la generación de texto, sino que también reproduce las desigualdades inherentes al sistema de conocimiento global.

Al respecto, Gallent-Torres et al. (2023) analizan cómo herramientas como ChatGPT representan una amenaza para la integridad académica, al ser capaces de generar textos que imitan la autoría humana con un nivel de sofisticación que dificulta su detección. Esta situación plantea interrogantes éticos sobre la originalidad y autenticidad del trabajo académico, así como sobre los criterios de evaluación en contextos educativos. La posibilidad de presentar textos generados por IA como propios socava los principios fundamentales de la educación, basados en el desarrollo del pensamiento crítico, la argumentación y la creatividad. Por ejemplo, un estudio de Alvero, Lee, Regla Vargas y otros (2024) mostró cómo los textos de ensayos generados con IA tienen más similitud a los escritos por estudiantes varones de altos recursos. Además, el uso indebido de estas tecnologías podría trivializar la producción académica, reduciendo el aprendizaje a un proceso mecánico de reproducción textual sin comprensión profunda.

De esta forma, la capacidad de la IAG para replicar patrones lingüísticos se basa en algoritmos que identifican correlaciones estadísticas en grandes volúmenes de datos. Sin embargo, esta habilidad técnica también conlleva la reproducción de sesgos presentes en los datos de entrenamiento. Por ejemplo, Pérez (2024) advierte que, en tareas de análisis de sentimientos, términos como “emocional” o “agresiva” se asocian desproporcionadamente con textos escritos por mujeres, lo que refuerza estereotipos de género. Además, las capacidades generativas competentes de los LLM solo se observan en idiomas con altos recursos, mientras que lenguas con menor cantidad de hablantes no gozan de las mismas capacidades debido al desequilibrio en los datos previos al entrenamiento (Nguyen et al, 2024). Este tipo de sesgos no son meros errores técnicos, sino reflejos de desigualdades estructurales que se perpetúan a través de los sistemas de IA. Además, la falta de representación de lenguas minoritarias en los datos de entrenamiento de modelos como GPT-4 profundiza la brecha lingüística y cultural, lo que lleva a la invisibilización de comunidades enteras en el ámbito digital.

Por lo tanto, la opacidad en los procesos de entrenamiento y desarrollo de modelos generativos también constituye una barrera para la evaluación crítica de su funcionamiento. Empresas como OpenAI no divulgan información completa sobre los

datasets utilizados, lo que impide una revisión independiente y transparente. Li et al. (2024) sostienen que esta falta de transparencia incrementa la posibilidad de que los modelos reproduzcan sesgos sin que existan mecanismos claros para su detección y corrección. Este problema se agrava cuando se considera que los usuarios tienden a asumir la neutralidad de la IA, desconociendo que los corpus textuales que alimentan estos modelos son construcciones sociales cargadas de ideología. La ausencia de una evaluación crítica de estas tecnologías pone en riesgo la producción y difusión de conocimiento equitativo y plural.

Dwork & Roth (2014), explica que las decisiones de diseño en los algoritmos incorporan juicios de valor y, por lo tanto, sesgan el modo en que funcionan los sistemas. Recordemos que la minería de datos aprende con el ejemplo, por lo cual depende de la información que recibe y ha estado expuesta y en el proceso de selección de características, las organizaciones toman decisiones sobre los atributos que observan y posteriormente incorporan en sus análisis.

Al respecto se encuentra cómo empresas de IA suelen defender sus modelos como "neutrales", argumentando que los sesgos reflejan "la realidad de los datos", pero esta postura ignora que los corpus textuales son construcciones sociales cargadas de ideología. Existen varios ejemplos en cuestión, como el PLN clínico donde lenguas minoritarias sufren inequidades por falta de datos (Névéol et al., 2018a), o el caso donde modelos como ChatGPT o Gemini reflejan desigualdades históricas (Kessler & Casal, 2024; Névéol et al., 2018a).

A esto se suma cómo herramientas como ChatGPT amenazan la integridad académica al generar textos que imitan la autoría humana (Gallent-Torres et al., 2023), entre muchos otros casos. El trabajo de Rentier (2024) muestra que la generación de textos puede carecer de precisión fáctica o lógica y que GenAI no es transparente frente a dónde se obtiene la información.

Frente a estos desafíos, se han propuesto estrategias para promover un uso ético e inclusivo de la IAG. Una de ellas es la diversificación de los datos de entrenamiento, integrando lenguas y perspectivas tradicionalmente marginadas. Muñoz-Basols et al. (2024) proponen el marco IMI+, que integra enfoques decoloniales en el diseño y entrenamiento de modelos de IA, exigiendo diversidad de fuentes, transparencia en el etiquetado de datos y participación de sociolingüistas en los procesos de evaluación. Sin embargo, estas iniciativas enfrentan obstáculos significativos debido a la falta de colaboración entre las comunidades académicas y las corporaciones tecnológicas, así como a la resistencia de estas últimas a revelar sus procesos internos.

Así es como la participación comunitaria emerge como una estrategia clave para contrarrestar las desigualdades estructurales en el desarrollo de IAG (Loftus et al, 2024) subraya la necesidad de involucrar a comunidades locales y hablantes nativos en los procesos de recolección y etiquetado de datos, garantizando una representación auténtica de sus voces. Esta colaboración no solo enriquecería la diversidad lingüística y cultural de los modelos generativos, sino que también contribuiría a su legitimidad y

aceptación en contextos locales. Además, la educación crítica en el uso de tecnologías de IA se presenta como una herramienta fundamental para empoderar a los usuarios y fomentar un uso consciente de estas herramientas.

En efecto, la integración de enfoques éticos y la gobernanza transparente son aspectos imprescindibles para garantizar que la IAG no perpetúe desigualdades estructurales. Kessler y Casal (2024) sugieren que los lingüistas y sociolingüistas deben desempeñar un papel central en la creación de marcos éticos para la IA, estableciendo principios de transparencia, equidad y responsabilidad. Asimismo, proponen la realización de auditorías externas independientes que permitan evaluar los sesgos presentes en los modelos de IA, asegurando que las decisiones algorítmicas sean justas y explicables. Estas medidas contribuirían a una regulación más estricta del desarrollo de IAG, promoviendo la inclusión y la diversidad lingüística.

La interdisciplinariedad también desempeña un papel crucial en el diseño de modelos de IAG más robustos y culturalmente sensibles. La colaboración entre lingüistas, ingenieros, neurocientíficos y expertos en ética puede conducir a la creación de modelos que emulen procesos cognitivos humanos, como la memoria episódica, la integración multisensorial y el razonamiento pragmático. Inoue (2024) propone modelos ecológicos de IA que incorporen variaciones dialectales y contextos culturales, priorizando la creación de modelos ligeros y accesibles para regiones con infraestructura tecnológica limitada. Esta aproximación no solo mejoraría la comprensión lingüística de los modelos generativos, sino que también garantizaría su relevancia y aplicabilidad en contextos diversos.

Y, aunque la inteligencia artificial generativa (IAG) promete ser una herramienta universal, pero su desarrollo está marcado por una paradoja: mientras aspira a trascender fronteras, reproduce jerarquías lingüísticas y cognitivas profundamente arraigadas. La brecha entre lenguas hegemónicas y minorizadas, así como la desconexión entre la cognición humana y la artificial, no son meros obstáculos técnicos, sino síntomas de un problema estructural.

Estos retos exigen una revisión crítica de cómo la IAG conceptualiza el lenguaje y la inteligencia, y cómo reproduce —o subvierte— desigualdades históricas. En efecto, los modelos de IAG se entrenan con datasets que reflejan un orden lingüístico global desigual. Por ejemplo, el 92% de los datos de GPT-4 provienen de fuentes en inglés, chino mandarín y español, mientras que lenguas como el quechua, el náhuatl o el yorúba representan menos del 0.3% (Bolaño-García, 2024). Esta desproporción no es casual: responde a dinámicas coloniales que privilegian idiomas asociados con poder económico y político. El resultado es una IAG que invisibiliza a más de 3,000 lenguas, acelerando su extinción y marginando a sus hablantes.

Aquí, la lingüística de corpus aporta metodologías para recolectar y etiquetar textos en lenguas de bajos recursos, como estrategia para democratizar la IAG, por un lado, se propone que los lingüistas lideren los marcos éticos para IAG, con el fin de promover transparencia en fuentes y mitigar estereotipos culturales (Kessler y Casal, 2024), y por

otro lado, Paulus y Marone (2024) advierten sobre cómo al integrar herramientas GenAI en software de codificación cualitativa para el análisis en estudios cualitativos como ATLAS.ti, NVivo y MAXQDA se puede entrar en conflicto con fundamentos epistemológicos propios de los estudios críticos del discurso.

Además, la lingüística crítica, desde enfoques como el análisis del discurso, proporciona herramientas para deconstruir estas dinámicas. Por ejemplo, el marco IMI+ de Muñoz-Basols et al. (2024) integra perspectivas decoloniales en el entrenamiento de IA, que exige: diversidad de fuentes, inclusión de textos escritos por comunidades marginadas, transparencia en etiquetado, documentación pública de criterios para identificar sesgos y participación de sociolingüistas en la evaluación de modelos. Sin embargo, estas iniciativas chocan con la opacidad corporativa. Empresas como OpenAI no revelan detalles completos de sus datasets, lo que dificulta una evaluación independiente (Li et al., 2024).

Es así como la lingüística no puede limitarse a ser un insumo técnico para la IAG; debe erigirse como un contrapoder ético. Los sesgos no son “errores” corregibles con ajustes algorítmicos, sino síntomas de estructuras de poder que privilegian ciertas voces sobre otras. Sin una intervención lingüística crítica, la IAG seguirá siendo un instrumento de homogenización cultural, lejos de su promesa de democratizar el conocimiento. A pesar de los avances, la IAG prioriza lenguas mayoritarias. Estudios como el de Jain et al. (2024) muestran que el acceso desigual a tecnologías digitales amplía brechas educativas, mientras que Inoue (2024) propone modelos ecológicos que integren variaciones dialectales y contextos culturales.

A lo anterior se suma el hecho de que la inteligencia artificial esté siendo adoptada cada vez más en el sector público, especialmente en la seguridad pública en donde la tecnología de identificación biométrica remota (RBI) es objeto de un intenso debate pues perpetúan sesgos y vulneran derechos humanos fundamentales (Kieslich & Lünich, 2024).

En primer lugar, se deben fortalecer los esfuerzos de documentación y digitalización de lenguas minoritarias, asegurando que estos datos sean accesibles para la comunidad científica y tecnológica. En segundo lugar, se requiere la implementación de políticas que obliguen a las grandes corporaciones a considerar la diversidad lingüística en el diseño y entrenamiento de sus modelos. Finalmente, la interdisciplinariedad entre lingüística, IA y ciencias cognitivas es clave para desarrollar modelos que no solo sean eficientes en la generación de texto, sino que también respeten la pluralidad lingüística y cultural. Frente a esta situación y a la compleja estructura de la IAG, el grupo Fairness, Accountability and Transparency (FAcCT) reúne la participación interdisciplinaria de profesionales que debaten los problemas de discriminación, falta de transparencia y rendición de cuentas, y ausencia de un debido proceso en los sistemas automatizados.

Por lo tanto, la diversidad lingüística es un pilar ético y técnico por lo que ignorarlo no solo perpetúa desigualdades, sino que condena a la IA a una comprensión superficial y

etnocéntrica del lenguaje. La verdadera innovación no radica en modelos más grandes, sino en paradigmas más humildes y colaborativos, donde la lingüística y las comunidades marginadas lideren el diseño tecnológico. Solo así la IAG podrá aspirar a ser verdaderamente universal, por lo que, la diversidad lingüística y la cognición representan desafíos fundamentales para el desarrollo de la IAG. La dependencia de lenguas mayoritarias, la falta de modelos pragmáticamente competentes y la colonialidad en la producción tecnológica limitan el potencial de estas herramientas para convertirse en verdaderos agentes comunicativos. La solución a estos problemas no solo depende de avances técnicos, sino también de un compromiso ético y político con la equidad lingüística y la representación justa en la era digital.

Finalmente, es fundamental abordar la colonialidad en el desarrollo de tecnologías lingüísticas. Bolaño-García (2024) advierte que los modelos de IAG están diseñados y entrenados mayoritariamente en el Norte Global, lo que implica que reflejan y refuerzan visiones del mundo dominantes, dejando de lado epistemologías y formas de conocimiento propias de comunidades no hegemónicas. Esta tendencia contribuye a la marginalización de lenguas indígenas y a la reproducción de sesgos culturales que distorsionan la representación de ciertas identidades en los sistemas de IA. Para contrarrestar este fenómeno, se requieren políticas públicas y legislaciones que promuevan la diversidad lingüística y cultural en el desarrollo de tecnologías de IA, así como la financiación de proyectos liderados por comunidades marginadas.

Una muestra de ello es el trabajo de Koc (2025) en donde analiza la ética y las limitaciones las IA generativas y los LLM en la preservación de lenguas en peligro de extinción y ve como oportunidad la capacidad de producir grandes volúmenes de materiales lingüísticos, así como apoyar la investigación lingüística identificando patrones y estructuras que no son muy evidentes, pero también alerta frente a posibles intereses económicos y la obtención de materiales lingüísticos de código abierto sin el consentimiento de la comunidad.

También se encuentra el trabajo de Uchechukwu (2025) que busca superar el problema de la memoria a largo plazo de las IAG mediante memoria episódica para el aprendizaje experiencial, memoria semántica para la comprensión conceptual y memoria procedimental para la retención de habilidades.

Conclusiones

En conclusión, aunque la inteligencia artificial generativa ofrece oportunidades significativas para democratizar el conocimiento y optimizar procesos cognitivos, su uso generalizado plantea una serie de desafíos éticos, sociales y lingüísticos que no pueden ser ignorados. La reproducción de sesgos lingüísticos y sociales, la amenaza a la integridad académica y la perpetuación de desigualdades estructurales son problemas que exigen un enfoque crítico y multidisciplinario. Solo mediante la colaboración entre comunidades locales, expertos en lingüística, ingenieros y responsables políticos se podrá garantizar el desarrollo de tecnologías de IAG que sean inclusivas, éticas y culturalmente sensibles. La verdadera innovación en este campo no radica únicamente

en la creación de modelos más grandes y eficientes, sino en la construcción de paradigmas tecnológicos que respeten y promuevan la pluralidad lingüística y cultural del mundo.

Retos y reflexión

Para promover una inteligencia artificial (IA) más inclusiva, ética y cognitivamente robusta, es necesario adoptar un enfoque interdisciplinario que integre conocimientos de diversas disciplinas, como la lingüística, la ciencia cognitiva, la ética, las ciencias sociales, la tecnología y las humanidades. Por ejemplo, hacer la documentación y digitalización de lenguas minoritarias donde se involucre a hablantes nativos y comunidades locales en la recolección y etiquetado de datos, garantizando que sus voces sean auténticamente representadas.

Diseñar modelos que no solo traduzcan palabras, sino que también capturen conceptos culturales únicos y cosmovisiones. Esto significa incorporar variaciones dialectales y registros sociolingüísticos para evitar la homogeneización cultural.

Descentralizar la producción de tecnologías de IA, involucrando a instituciones académicas y tecnológicas del Sur Global en el diseño y entrenamiento de modelos.

Crear programas educativos que combinen IA, lingüística, ética y ciencias sociales, formando profesionales capaces de abordar los desafíos complejos de la IA desde múltiples perspectivas en donde se capacite a docentes y estudiantes en el uso crítico de herramientas de IA, enseñándoles a identificar y mitigar sesgos.

Referencias bibliográficas

- Alvero, A. J., Lee, J., Regla-Vargas, A. et al. (2024). *Large language models, social demography, and hegemony: comparing authorship in human and synthetic text*. <https://doi.org/10.1186/s40537-024-00986-7>
- Beerends, S. & Aydin, C. (2024). *Negotiating the authenticity of AI: how the discourse on AI rejects human indeterminacy*. <https://doi.org/10.1007/s00146-024-01884-5>
- Bernal Leongómez, J. (1984). *Tres momentos estelares en lingüística*. <https://selloeditorial.caroycuervo.gov.co/detalle-libro/tres-momentos-estelares-en-linguistica-pdf>
- Berwick, R. C., Chomsky, N. et al. (2013). *Evolution, brain, and the nature of language*. <https://doi.org/10.1016/j.tics.2012.12.002>
- Bolaño García, M. (2024). Inteligencia artificial para la educación: desafíos y oportunidades. *Praxis*, 20(1), 8–12. <https://doi.org/10.21676/23897856.5997>
- Brown, T. et al. (2020). *Language Models are Few-Shot Learners*. <https://arxiv.org/abs/2005.14165>
- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

- Colin, R. et al. (2020). *Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer*. <https://jmlr.org/papers/volume21/20-074/20-074.pdf>
- Devlin, J., Chang, M. W., Lee, K. & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding*. *Proceedings of NAACL-HLT 2019*. <https://www.aclweb.org/anthology/N19-1423/>
- Dwork, C. & Roth, A. (2014). The Algorithmic Foundations of Differential Privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4), 211-407. <http://dx.doi.org/10.1561/0400000042>
- Fairclough, N. (1992). *Discourse and social change*. Polity Press.
- Gallent-Torres, C. et al. (2023). *El impacto de la inteligencia artificial generativa en educación superior: una mirada desde la ética y la integridad académica*. <https://doi.org/10.30827/relieve.v29i2.29134>
- Greco, C. M. & Tagarelli, A. (2024). Ordenando el ámbito de los modelos lingüísticos basados en Transformers para la inteligencia artificial y el derecho. *Artif Intell Law*, 32, 863–1010. <https://doi.org/10.1007/s10506-023-09374-7>
- Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. En Joseph H. Greenberg (ed.), *Universals of Language* (2ª ed.), 73-113. MIT Press. <https://babel.ucsc.edu/~hank/114/greenberg.pdf>
- Inoue, H. (2024). Ecological AI models: Towards linguistic diversity and inclusivity. *Linguistic Frontiers in AI Research*, 21(4), 56-72. https://www.researchgate.net/publication/385652682_The_Education_Revolution_through_Artificial_Intelligence_Enhancing_Skills_Safeguarding_Rights_and_Facilitating_Human-Machine_Collaboration
- Jain, R. et al. (2024). Digital divide and AI accessibility: An educational perspective. *Computers in Education Review*, 32(1), 100-118. https://www.researchgate.net/publication/371575436_Addressing_the_Digital_Divide_Access_and_Use_of_Technology_in_Education
- Jackson, C. (2018). *Learn programming in python with Cody Jackson: Grasp the basics of programming and Python syntax while building real-world applications*. <https://www.abebooks.co.uk/9781789531947/Learn-Programming-Python-Cody-Jackson-1789531942/plp>

- Kessler, T. & Casal, M. (2024). The role of linguists in AI ethics: Transparency and bias mitigation. *Journal of Applied Linguistics and AI Ethics*, 18(2), 65-84. https://www.researchgate.net/profile/George-Benneh-Mensah/publication/375744287_Artificial_Intelligence_and_Ethics_A_Comprehensive_Review_of_Bias_Mitigation_Transparency_and_Accountability_in_AI
- Kieslich, K. & Lünich, M. (2024). *Regulating AI-Based Remote Biometric Identification. Investigating the Public Demand for Bans, Audits, and Public Database Registrations.* <https://facctconference.org/static/papers24/facct24-12.pdf>
- Koc, V. (2025). *Generative AI and Large Language Models in Language Preservation: Opportunities and Challenges.* <https://arxiv.org/html/2501.11496v1>
- Li, B., Nye, M. & Andreas, J. (2021). *Implicit Representations of Meaning in Neural Language Models.* In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. <https://doi.org/10.18653/v1/2021.acl-long.143>
- Li, X. et al. (2024). *The transparency challenge in AI: Dataset opacity and accountability.* <https://doi.org/10.1016/j.dss.2020.113302>
- Mohammed, Q. & Shormani, I. (2024). *Generative linguistics contribution to artificial intelligence: Where this contribution lies?* <https://www.arxiv.org/pdf/2410.20221>
- Muñoz Basols, J., Palomares Marín, M. & Moreno Fernández, F. (2024). *El Sesgo Lingüístico Digital (SLD) en la inteligencia artificial: implicaciones para los modelos de lenguaje masivos en español.* <https://doi.org/10.15381/lengsoc.v23i2.28665>
- Nguyen, A. et al. (2023). Democratizing AI: The role of linguistic corpora in low-resource languages. *Journal of Computational Linguistics and AI*, 45(3), 125-140. <https://doi.org/10.48550/arXiv.2306.11372>
- Névéol, A., Dalianis, H., Velupillai, S., Savova, G. & Zweigenbaum, P. (2018a). Clinical Natural Language Processing in languages other than English: opportunities and challenges. *J Biomed Semantics*, 9(1), 12. <https://jbiomedsem.biomedcentral.com/articles/10.1186/s13326-018-0179-8>
- Névéol, A., et al. (2018b). Clinical NLP for underrepresented languages: Challenges and perspectives. *Journal of Biomedical Informatics*, 88, 125-142. <https://jbiomedsem.biomedcentral.com/articles/10.1186/s13326-018-0179-8>
- Paulus, T. M. & Marone, V. (2024). *In minutes instead of weeks: Discursive constructions of generative AI and qualitative data analysis.* <https://doi.org/10.1177/10778004241250065>

- Pannachin, C. & Goertzel, B. (2007). Contemporary approaches to artificial general intelligence. ennachin, C., Goertzel, B. (2007). Contemporary Approaches to Artificial General Intelligence. In Goertzel, B., Pennachin, C. (eds) *Artificial General Intelligence. Cognitive Technologies*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-68677-4_1
- Pérez Ugena, C. (2024). Sesgo de género (en IA). *Eunomía. Revista en Cultura de la Legalidad*, 26, 311-330. <https://doi.org/10.20318/eunomia.2024.8515>
- Rentier, E. (2024). *To use or not to use: exploring the ethical implications of using generative AI in academic writing*. <https://doi.org/10.1007/s43681-024-00649-6>
- Riccardi, A., et al. (2024). The Two Word Test: Evaluating semantic coherence in large language models. *Computational Semantics Review*. <https://www.nature.com/articles/s41598-024-72528-3.pdf>
- Loftus, T. J., Balch, J. A., Abbott, K. L., Hu, D., Ruppert, M. M., Shickel, B., Ozrazgat-Baslanti, T., Efron, P. A., Tighe, P. J., Hogan, W. R., Rashidi, P., Cardel, M. I. Upchurch, G. R. & Bihorac A. (2024). *Community-engaged artificial intelligence research: A scoping review*. <https://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000561>
- Uchekukwu, A. (2025). *Memory Architectures in Long-Term AI Agents: Beyond Simple State Representation*. <https://doi.org/10.13140/RG.2.2.26486.51527>
- Van Dijk, T. A. (1993). *Elite discourse and racism*. SAGE Publications.
- Watson, D. (2019). The Rhetoric and Reality of Anthropomorphism in Artificial Intelligence. *Minds & Machines* 29, 417–440. <https://doi.org/10.1007/s11023-019-09506-6>
- Zhang, Y. et al. (2023). *SemBERT: Enhancing BERT with semantic role labeling*. <https://ojs.aaai.org/index.php/AAAI/article/view/6510>

Conflicto de intereses: Los autores declaran no tener conflictos de intereses.

Contribución de autores: Los autores participaron en la búsqueda y análisis de la información para el artículo, así como en su diseño y redacción.