

# Inteligencia artificial y comprensión. Sobre la artificial manera de considerar la comprensión humana

Rudyard Loyola Cortés  
Universidad de Chile 

<https://dx.doi.org/10.5209/ref.96220>

Recibido: 28/05/2024 • Aceptado: 24/03/2025 • Publicado en línea: 16/06/2025

**Resumen:** A través de diversas contribuciones filosóficas – el mundo (*Welt*) y el uno (*das Man*) heideggerianos, la percepción de Merleau-Ponty, el trasfondo de Hubert Dreyfus y Charles Taylor y la *Lebensform* de Wittgenstein – se busca contrastar un esquema fenomenológico-ontológico de comprensión-interpretación con el modelo tradicional, de índole epistemológico-antropológico, que considera al ser humano un sujeto cognoscente. A partir de ello, se compara dos paradigmas de inteligencia Artificial (IA) – los modelos simbólico y estadístico neural – con estos dos modos de entender el afrontamiento con la realidad, siguiendo la reflexión de Hubert Dreyfus. Finalmente, se presenta el paradigma de trabajo con los patrones del lenguaje de los programas *LLM* como una alternativa diversa, dentro del paradigma estadístico neural, recurriendo a una concepción del lenguaje del Heidegger posterior a *Ser y tiempo*.

**Palabras Clave:** inteligencia artificial (IA); comprensión; trasfondo; mundo (*Welt*); modelo simbólico; modelo neural; Dreyfus; Taylor; Heidegger.

## EN Artificial Intelligence and Understanding. On the Artificial Way of Considering Human Understanding

**Abstract:** Through various philosophical contributions—the Heideggerian concepts of the world (*Welt*) and the they (*das Man*), Merleau-Ponty's perception, the background of Hubert Dreyfus and Charles Taylor, and Wittgenstein's *Lebensform*—this work aims to contrast a phenomenological-ontological framework of understanding-interpretation with the traditional epistemological-anthropological model, which considers the human being as a knowing subject. Based on this, two paradigms of Artificial Intelligence (AI)—the symbolic model and the neural statistical model—are compared with these two ways of understanding engagement with reality, following the reflections of Hubert Dreyfus. Finally, the paradigm of working with the language patterns of LLM (Large Language Model) programs is presented as a diverse alternative within the neural statistical paradigm, drawing on a conception of language from the later Heidegger, after *Being and Time*.

**Keywords:** artificial intelligence (AI); understanding; background; world (*Welt*); symbolic model; neural model; Dreyfus; Taylor; Heidegger.

**Sumario:** 1. Introducción; 2. El mundo de Heidegger y la comprensión; 3. La percepción: Taylor, Merleau-Ponty y Heidegger; 4. El trasfondo de Dreyfus & Taylor y el uno heideggeriano; 5. Wittgenstein y la *Lebensform*; 6. El modelo computacional de la mente; 7. Modelos simbólico y neural de la IA; a. El modelo simbólico y su trasfondo filosófico; b. El modelo estadístico-neural; c. El problema del sentido común en la reflexión de Dreyfus sobre la IA; 8. Los LLM y la interacción con el ser humano; 9. Conclusión; 10. Referencias bibliográficas.

**Cómo citar:** Loyola Cortés, R. "Inteligencia artificial y comprensión. Sobre la artificial manera de considerar la comprensión humana", *Revista de Filosofía*, avance en línea, <https://dx.doi.org/10.5209/ref.96220>

## 1. Introducción

La manera en que hemos interpretado el conocimiento humano ha abierto la posibilidad de compararnos de manera no suficientemente reflexiva con el computador (ordenador), trasponiendo su funcionamiento a nuestra propia forma de comprender. En la manera de entender cómo conocemos, la filosofía ha tenido un rol relevante que ha influido mucho más allá de los ámbitos disciplinares donde se desarrolla. Esto nos lleva a preguntarnos cómo la reflexión filosófica, específicamente sobre el conocimiento, ha incidido en lo que llamamos “inteligencia artificial” (IA) y cómo de manera inversa la concepción que tenemos de la IA ha servido para entender nuestra propia comprensión humana.

Un tipo de modelo de IA, pretendiendo simular la inteligencia humana, ha tomado como paradigma una tradición epistemológica que entiende el conocimiento como una relación entre un sujeto y un objeto en la que median, ideas simples, sensaciones o impresiones sensibles. Autores contemporáneos han criticado esta manera de entender el conocimiento y la consideración antropológica que le subyace. Al respecto, nos proponemos reflexionar sobre diversas contribuciones filosóficas, que recoge el filósofo norteamericano Hubert Dreyfus y el canadiense Charles Taylor, desde la reflexión sobre su noción de trasfondo (*Background*). Para tal efecto, reuniremos, en primer lugar, algunos aportes filosóficos: las reflexiones de Heidegger (*Ser y tiempo*) sobre su forma de entender la comprensión humana, algunas reflexiones de Merleau-Ponty sobre la percepción y de Wittgenstein en torno a la inteligibilidad. Tras lo ganado en estas reflexiones – y siguiendo los análisis de Hubert Dreyfus – pretendemos comparar estas formas de entender la comprensión humana con los dos principales modelos de IA: el modelo simbólico y el modelo estadístico-neural. En definitiva, vamos a contraponer un modelo de inteligencia –que sirvió de base para una simulación de IA de acuerdo con el paradigma simbólico – con un modelo comprensivo-interpretativo, que ya no se centra tanto en la diada inteligencia-conocimiento, sino en la diada comprensión-praxis y que pensamos se acerca al paradigma estadístico neural de la IA. Como podremos constatar el esquema comprensivo-interpretativo requiere de un elemento no considerado en la manera tradicional de entender el conocimiento: un contexto de significatividad, que con Heidegger podemos llamar “mundo” (*Welt*) y con Dreyfus y Taylor “trasfondo” (*Background*).

Además, revisaremos algunos logros actuales del modelo estadístico neural que entregan nuevas perspectivas, en particular, el modo en que funcionan los programas *LLM* (*Large Language Model*), que se han hecho conocidos en programas como Chat GPT, Gemini, Llama y Claude. De igual modo, trataremos de interpretarlos desde un modelo de comprensión humana que integre el trasfondo de significatividad.

## 2. El mundo de Heidegger y la comprensión

Comenzamos con las reflexiones de Martin Heidegger sobre su crítica al esquema moderno del conocimiento del mundo. El filósofo alemán critica el esquema sujeto-objeto y en especial la noción de sujeto. Este se comprende como un ser encerrado en sí mismo que tiene que trascender su propia interioridad para acceder al más allá del mundo y poder conocer al objeto. Así las cosas, el conocimiento es entendido como una correlación entre un sujeto y un objeto (Cfr. Heidegger, 1998, p. 86, Cfr. 1927, p. 60; Cfr. Hessen, 2006, p. 30)<sup>1</sup>. En este esquema hay una separación radical entre la interioridad humana y la exterioridad del mundo. Los grandes problemas filosóficos sobre el conocimiento surgen principalmente de este asunto: la incomunicabilidad entre el interior del ser humano y la exterioridad del mundo, ¿cómo es posible trascender la inmanencia del sujeto para acceder al mundo donde se encuentra el objeto y desde ahí capturar una imagen de él? La respuesta de Heidegger es que este esquema sujeto-objeto es una interpretación externa y formal (Cfr. Heidegger, 1998, p. 86, Cfr. 1927, p. 60), creadora de falsos problemas.

Si bien sujeto y objeto son dos conceptos que sirven para el análisis del conocimiento, en la realidad no hay, para Heidegger, algo así como un sujeto que tenga que trascenderse a sí mismo para acceder al mundo y encontrarse en él con el objeto. Esto porque en la concepción heideggeriana el *Dasein* – que en este trabajo entenderemos como ser humano – es un ser-en-el-mundo. No tiene que trascenderse para alcanzar el mundo, porque el mundo es un momento de su ser. Estas afirmaciones están basadas en un análisis fenomenológico que parte de la comparecencia de las cosas en la cotidianidad. Heidegger advierte que el *Dasein*, la mayor parte de las veces no es un sujeto cognosciente, sino alguien absorto en la praxis de la cotidianidad. En ella el ente que le comparece inmediatamente es el útil [*Das Zeug*] (Cfr. Heidegger, 1998, p. 96, Cfr. 1927, p. 68). Este es definido como el ente que está siendo usado o producido en la praxis cotidiana (Cfr. Heidegger, 1998, p. 95, Cfr. 1927, p. 67). El útil no es un objeto de conocimiento, es opaco, explícitamente invisible. Un útil como un martillo en el martillar no comparece explícitamente ante nosotros. Absortos en lo que hacemos, nos comparece de manera inadvertida pues estamos preocupados

<sup>1</sup> Heidegger en el §13 de *Ser y tiempo* – “Ejemplificación del estar-en por medio del modo fundado. El conocimiento del mundo” – critica la manera de entender el conocimiento de Nicolai Hartmann. Johannes Hessen expresa fundamentalmente esta línea de interpretación (Cfr. Acevedo, 1999, p. 33).

de aquello que estamos haciendo. Dice Heidegger del útil que tiene el modo de ser de estar-a-la-mano [*Zuhandensein*] (Cfr. Heidegger, 1998, p. 97, Cfr. 1927, p. 69).

En un análisis más profundo, el útil nos aparece conectado con otros útiles: Vgr. el martillo, remite a un clavo, este a una tabla, esta a lo que se está construyendo, digamos una mesa. Tales relaciones aparecen en la comparecencia que tenemos de las cosas en la praxis de la cotidianidad. Las cosas en ella están vinculadas por unas relaciones llamadas «remisiones» (Cfr. Heidegger, 1998, p. 96, Cfr. 1927, p. 68). De todas ellas hay algunas que remiten al *Dasein*. Estas remisiones forman un plexo que nos otorga significatividad (Cfr. Heidegger, 1998, p. 113, Cfr. 1927, p. 87) y gracias a este entramado podemos conocer las cosas como lo que son en la cotidianidad. Es a este plexo de significatividad que Heidegger llama «el mundo circundante» y es el que permite la comparecencia de las cosas. Desde la significatividad que otorga este plexo, que constituye el mundo circundante, las cosas se nos aparecen (Cfr. Heidegger, 1998, p. 214, Cfr. 1927, p. 192). No son los entes los que constituyen el mundo – en tanto momento de ser-en-el-mundo – sino que son estas relaciones llamadas «remisiones». Ellas nos otorgan la significatividad que posibilita la comparecencia de las cosas. Un cambio de mundo o contexto podría ocasionar que las cosas nos comparezcan de manera distinta. Por ejemplo, un ordenador (computador) es lo que es en esta sociedad en que la tecnología juega un rol relevante; en una sociedad prehistórica este artefacto no comparecería de la misma forma, pudiendo incluso ser considerado sagrado o maligno.

El *Dasein* es un ser que se encuentra ocupado con los útiles en la cotidianidad (Cfr. Heidegger, 1998, p. 354, Cfr. 1927, p. 337). Es desde este modo de ser originario que puede acceder al modo de ser del conocimiento (Cfr. Heidegger, 1998, p. 86 ss., Cfr. 1927, p. 59 ss.). Aquí se traslucen una opción metodológica investigativa distinta: la elección del modo de acceso al ser humano con respecto a la tradición filosófica (Cfr. Heidegger, 1998, pp. 40-41, Cfr. 1927, p. 16). Para esta última, el modo de ser por antonomasia del ser humano es el conocimiento y, por ende, su modo de acceso obvio. Para Heidegger, en cambio, el modo en que el *Dasein* está la mayor parte de las veces es la praxis de la cotidianidad. Es desde aquí, que, en ocasiones, el *Dasein* puede acceder a ser un ser-en-el-mundo cognosciente (Cfr. Heidegger, 1998, p. 86, Cfr. 1927, p. 60). En este cambio se opera una deficiencia tanto en el *Dasein* como en el útil. Para que este último comparezca como objeto – ente que aparece en el modo de ser del conocimiento – debe interrumpirse la praxis de la cotidianidad y el útil perder sus remisiones. Es decir, desmundanizarse o descontextualizarse (Cfr. Heidegger, 1998, p. 102, Cfr. 1927, p. 75). Es desde la praxis de la cotidianidad que accedemos al modo de ser del conocimiento. Este último es menos originario y menos frecuente (Vattimo, 2006, p. 25).

El *Dasein* es siempre un ser-en-el-mundo, todas sus posibilidades de existencia, requieren del mundo en tanto momento de su ser. El norteamericano Hubert Dreyfus, ilustra cómo esto ocurre desde la más temprana edad, incluso antes de que poseamos lenguaje: compara dos bebés, uno norteamericano y otro japonés. Desde la más temprana infancia ambos están insertos en mundos distintos. El bebé norteamericano se cría en un contexto que privilegia la actividad y la espontaneidad, en él los niños son estimulados para que reaccionen ante las provocaciones que los adultos tiernamente les hacen. En cambio, el contexto japonés valora más el silencio. La madre acurruga más y no son tan comunes las provocaciones de adultos para que el bebé reaccione activamente. Como vemos, el mundo que rodea a cada niño es distinto y eso explica por qué los bebés también lo son (Cfr. Dreyfus, 1996, p. 161).

En la praxis de la cotidianidad, para Heidegger existe un tipo de «conocimiento» (con comillas) no teórico ni explícito: la circunspección [*Umsicht*] (Cfr. Heidegger, 1998, p. 97, Cfr. 1927, p. 69). Este es un «conocimiento» que nos permite movernos en el mundo. Cuando llegamos a una sala de clases, buscamos una silla y nos sentamos, lo hacemos gracias a la circunspección, ella nos guía en la praxis, sabemos en todo momento cómo movernos y qué hacer. Esto lo hacemos sin fijarnos explícitamente, la silla nos aparece de manera más bien opaca. Este «conocimiento» requiere de una comprensión aún más general que permite la interpretación en la situación fáctica: interpretamos el ente silla al sentarnos en ella. Absortos en lo que hacemos estamos involucrados con las cosas y no pensando explícitamente en ellas, no obstante, nos movemos con fluidez gracias a la circunspección. En ella nos movemos en la comprensión del ser de los entes y desde ella podemos interpretarlos en el uso que de ellos hacemos. En esto se juega también la interpretación de lo que nosotros somos (Heidegger, 1927, p. 12, Cfr. 1998, p. 35). Por ejemplo, nos entendemos como profesores y estudiantes. Esto gracias a la familiaridad que tenemos con el mundo y con sus remisiones (Heidegger, 1927, p. 86, Cfr. 1998, p. 113). Es desde la praxis de la cotidianidad que podemos acceder al conocimiento sin comillas, al conocimiento explícito. Es conveniente señalar que cuando estamos en el modo de ser del conocimiento también interpretamos, pero este tipo de interpretación es menos originaria (Cfr. Heidegger, 1927, p. 69, Cfr. 1998, p. 97). Esto debido a que es en el modo de ser de la praxis en el que nos encontramos por defecto y es desde este que podemos acceder al modo de ser del conocimiento<sup>2</sup>.

<sup>2</sup> Sobre esto trata el §13 de *Ser y tiempo*.

### 3. La percepción: Taylor, Merleau-Ponty y Heidegger

Charles Taylor analiza nuestro estar comprometidos con el mundo, basándose en la reflexión de Maurice Merleau-Ponty. Somos fundamentalmente sujetos encarnados y agentes corpóreos, que comprometidos con el mundo tratamos con realidades que son significativas para nosotros. Merleau-Ponty muestra esto a través del estudio fenomenológico de la percepción. Esta es nuestra vivencia más básica en el mundo. Todo lo demás como andar por la calle, conversar, trabajar, etcétera (Cfr. Taylor, 1989, p. 3) supone el acceso básico que se da a través de la percepción. Merleau-Ponty no considera esta última como impresiones sensibles, entendidas como átomos desprovistos de toda subjetividad. Esto último es incoherente pues el “campo de los fenómenos no es un ensamblaje de datos de los sentidos. Es nuestro acceso al mundo” (Taylor, 1989, p. 12). La percepción no se vive como captación de “datos de los sentidos”, sino que ya posee significatividad. Todo dato ya es una elaboración, una parcialidad propia de la selección de un agente encarnado inmerso en coordenadas espacio-temporales (Cfr. Llamas, 2001, p. 52). Nuestra percepción ya posee una estructura orientacional: Nuestro campo está compuesto de un primer plano y un fondo, un arriba y un abajo en relación con nuestra acción, vale decir, con nuestra capacidad de mantenernos en equilibrio y actuar (Cfr. Taylor, 1997, p. 47). Arriba y abajo, por ejemplo, no tiene que ver simplemente con cómo y dónde se encuentra nuestro cuerpo, tampoco con la ubicación de la cabeza y los pies, del cielo y la tierra – pensemos en un astronauta en el espacio sin las referencias que da la tierra y el cielo atmosférico – sino que abajo y arriba “están relacionados con cómo me movería y actuaría en el campo...” (Cfr. Merleau-Ponty, 1993, p. 265; Taylor, 1989, pp. 4-5, 1997, p. 47).

La articulación que da el campo es propia de un agente encarnado (Cfr. Taylor, 1997, p. 49) o de un *être-au-monde* (ser-del-mundo). Este campo se articula en zonas de accesibilidad e inaccesibilidad, de seguridad y peligro, por ejemplo, que tienen que ver con nuestra praxis. Así las cosas, el mundo y los entes intramundanos comparecen en función de lo que nos interesa, no siguiendo propiedades objetivas del entorno (Cfr. Taylor, 1989, p. 5). Esto dista de lo que ha señalado la epistemología moderna, sobre todo cartesiana y empirista, que ha considerado la percepción como un cuadro neutro. Somos en el mundo, al punto que este último «penetra hasta mi esencia» (Taylor, 1989, p. 8).

Por otro lado, percibir lo podemos entender como algo que hacemos y, a la vez, como algo que nos sucede. La percepción necesita también un sentido de fondo de la realidad circundante. Nuestra percepción opera en un mundo. El mundo de Merleau-Ponty lo entienden Dreyfus y Taylor como un trasfondo “que permite dar sentido a lo que nos rodea en términos de figura y fondo, ámbitos de restricción y de apertura, obstáculos y ayudas. En general se trata de un trasfondo que es compartido por todos los seres humanos y que se encuentra articulado en torno a lo que hemos llamado significados vitales” (Dreyfus & Taylor, 2016, p. 183). Incluso en una experiencia tan básica como nuestra percepción hay un trasfondo de significatividad.

### 4. El trasfondo de Dreyfus & Taylor y el uno heideggeriano

Volvamos al mundo heideggeriano, este es un contexto de significatividad compartido. Las cosas, o mejor, los entes nos comparecen desde un mundo esencialmente compartido [*Mitwelt*] (Cfr. Heidegger, 1998, p. 144, Cfr. 1927, p. 118). Para Dreyfus, el trasfondo se nos transmite fundamentalmente en la cotidianidad. Al respecto, Heidegger se pregunta por el quién de ella (Cfr. Heidegger, 1998, p. 140, Cfr. 1927, pp. 114-115). ¿Quién es el «sujeto» de la cotidianidad? No es tan fácil señalarlo, pues no es el *Dasein* individual, sino que son todos y a la vez nadie. La cotidianidad tiene como quién un sujeto impersonal. Ortega y Gasset se refería a él como «la gente», «la gente dice», «la gente hace». Este sujeto impersonal de la existencia, Heidegger lo denomina «*das Man*» (el uno). Insertos en la cotidianidad no somos nosotros los que vivimos, sino que es de alguna manera el uno el que nos vive, nos dice lo que debemos pensar, decir y hacer. Inmersos en la cotidianidad nos movemos en la predecible y segura vida de lo que está prefijado y que es obvio (Cfr. Heidegger, 2006, p. 352). El uno nos transmite prácticas sociales que nos permiten movernos con soltura en el mundo. Sabemos cómo saludar, qué decir y qué no, cómo comportarnos y cómo no. Para Heidegger en el uno nos encontramos en un modo inauténtico de ser. Desde él podemos acceder a modos más auténticos, pero es el uno el modo en que nos encontramos por defecto. Absortos en la cotidianidad nos movemos en la «tranquilidad» de lo familiar. Lo que da forma al uno no son reglas explícitas, sino en su mayoría directrices implícitas en las prácticas con las que nos desenvolvemos en él. Estas prácticas Dreyfus las llama «prácticas de trasfondo» [*background practices*] (Cfr. Dreyfus, 1996, p. 173, 2017, p. 184). Para Dreyfus, el uno es el existencial que nos permite la trasmisión del trasfondo. Vale decir, del contexto de significatividad. Además, detrás de nuestra comprensión no hay elementos últimos

<sup>3</sup> Jorge Eduardo Rivera, traductor chileno de *Ser y tiempo* al castellano, realiza un interesante comentario al respecto (En Heidegger, 1998, p. 474).

definidos, sino la ordinaria opacidad del uno. Una certeza última que sirva de base a nuestra inteligibilidad no existe (Cfr. Heidegger, 1997, p. 13; Cfr. Wittgenstein, 2009b nº 172).

El uno es una forma inauténtica de ser, con frecuencia se hace notar esto destacando su trivialidad, su repetitiva y poco original manera de existir. Pero el uno no solo tiene aspectos negativos, pertenece a la «reestructura positiva del *Dasein*» (Heidegger, 1998, p. 153, 1927, p. 129), es el existencial base del entendimiento cotidiano (Cfr. Dreyfus, 1996, p. 171). Esta publicidad del uno, dice Heidegger: «regula primeramente toda interpretación del mundo y del *Dasein*, y tiene en todo razón» (Heidegger, 1998, p. 151, 1927, p. 127). Desde esta comprensión previa e inauténtica del uno podemos acceder a otras formas de ser más auténticas, pero lo hacemos desde ella. Al decir Heidegger que el uno tiene «en todo razón», no está haciendo una afirmación veritativa lógica, sino que está estableciendo que al movernos en el uno nos movemos sin cuestionar, bajo sus directrices. Además, para Dreyfus es gracias al uno que el mundo público es compartido y transmitido (Dreyfus, 1996, p. 171; Cfr. Heidegger, 2006, p. 308).

## 5. Wittgenstein y la *Lebensform*

La reflexión de Dreyfus sobre el uno tiene además un componente wittgensteniano. En *Investigaciones Filosóficas*, el filósofo austriaco combate la concepción que entiende que una palabra adquiere su significado por la vinculación con un objeto. Esto es, un atomismo del significado. En un sentido análogo, John Locke considera que una palabra, si bien no se conecta con la cosa, sí lo hace con la idea mental que de ella se tiene. En esta senda podemos retrotraernos a otro tipo de atomismo, el atomismo de la mente: nuestros pensamientos podrían descomponerse en elementos últimos, en el caso de Descartes en ideas, ya sea innatas o adventicias, en el caso de Locke ideas simples, en el caso de Hume impresiones sensibles. Todos estos son elementos últimos que componen nuestro pensamiento a través de relaciones entre ellos. Esta concepción es simple y actualmente muy de sentido común, *inputs* que son procesados por operaciones formales. Vale decir, un tipo de modelo computacional: *inputs* – que pueden ser descompuestos en *bits* de información – que luego son procesados y que evacuan un *output*.

En el caso del atomismo que critica Wittgenstein, el problema estaría dado por la asociación simple entre palabra y cosa, conexión que le daría a la primera su significatividad. Para este filósofo, en cambio, la significatividad de las palabras estaría basada en juegos de lenguajes – por ejemplo, dar y recibir órdenes, relatar algo, inventar y contar historias, cantar, contar chistes, rezar, saludar y un amplio etcétera (Cfr. Marrades, 2014, p. 142) – los que a su vez están basados en formas de vida (*Lebensformen*), es decir, en formas de actuar. Este descubrimiento no es menor, nuestra comprensión o inteligibilidad no está basada en elementos últimos claramente definidos – sean estos ideas, impresiones sensibles, etcétera – sino en formas de actuar. Al fondo de nuestro pensar se encuentra una acción, no una idea o una impresión sensible. Es posible analizar la articulación de nuestros discursos y pensamientos, pero esto tiene un límite, llegados a lo más fundamental nos quedamos con un mero: «Así simplemente es como actúo» (Wittgenstein, 2009a, p. 339 Nº 217). Por tanto, el trasfondo no es totalmente articulable ni explicitable.

Para Taylor, tanto Heidegger, como Merleau-Ponty y Wittgenstein no entienden al ser humano y su comprensión como un lugar de representaciones, sino como un agente implicado en prácticas compartidas, tanto es así que son posibles las prácticas dialógicas, v.gr., el baile de una pareja de bailarines. Detrás de nuestra inteligibilidad y comprensión hay prácticas. Ya Heidegger se da cuenta de ello en sus análisis del útil y el mundo. Wittgenstein apunta a un trasfondo no basado en elementos atómicos definidos, sino en prácticas. Dreyfus concluye que tanto para Wittgenstein como para Heidegger «el origen de la inteligibilidad del mundo son las prácticas públicas promedio, las únicas mediante las cuales se puede llegar a una comprensión.» (Dreyfus, 1996, p. 172). No hay una certeza última basada en elementos claros y definidos en nuestra comprensión, sino prácticas compartidas.

## 6. El modelo computacional de la mente

Hasta aquí hemos visto dos modelos de afrontamiento humano con la realidad: uno que entiende al ser humano como un ser fundamentalmente cognosciente que genera pensamientos a partir de entradas (*inputs*) que pueden consistir en ideas simples, impresiones sensibles o elementos similares y, otro modelo, en que en nuestro acercamiento a la realidad media una comprensión de tipo holístico. En ella interpretamos desde un trasfondo de significatividad que nos precede y que es compartido (*Mitwelt*). Lo holístico de esta comprensión consiste en que las cosas nos comparecen de manera relacionada desde el trasfondo de significatividad. La comparecencia de algo depende de un ámbito más amplio. Vgr., un martillo nos comparece como tal desde el mundo circundante del taller. El primero de estos modelos entiende al ser humano como un ente encapsulado que recibe estímulos que luego procesa. El segundo modelo entiende que no hay estímulos que se consideren como elementos básicos a ser procesados. Incluso la percepción, tal como la entiende Merleau-Ponty, muestra que un estímulo perceptivo se nos aparece en relación con el mundo.

Es el primer modelo – el que entiende al ser humano como un sujeto cognoscente encerrado en sí mismo que recibe estímulos del exterior que luego procesa – al que Taylor y Dreyfus llaman el “modelo computacional de la mente”. Este es un paradigma que ha influido en cómo entendemos la comprensión humana. Paradójicamente este modelo se basa en una interpretación epistemológica de la mente humana que luego influyó en la comprensión de la IA. En esta dialéctica interpretativa confluyeron una serie de autores y tradiciones. Respecto a esto último, Taylor y Dreyfus ven algunas constantes que integradas generan lo que ellos llaman «el modelo mediacional del conocimiento». Para este paradigma no existiría un contacto directo con la realidad, sino que el conocimiento se daría mediado. En la tradición filosófica este modelo cuenta entre sus iniciadores a Descartes y a Locke. El primero partiendo de un escepticismo metódico, cree poder lograr la certeza al encontrar en su interior ideas claras y distintas, que a su vez sean indubitables. Las ideas son para Descartes entes mentales que median entre nosotros y la realidad (Cfr. Dreyfus & Taylor, 2016, pp. 25-32).

Al negar Locke la existencia de las ideas innatas cartesianas no niega el concepto de idea mental, solo que afirma que esta proviene de la experiencia: La realidad extramental comienza con una impresión en la mente, esto ocurre «por la operación de partículas insensibles sobre nuestros sentidos» (Locke, 1999, p. 114). Ocurre pasivamente, exenta de interpretación, al menos en su estado inicial de «idea simple». Este es un elemento puro y definido. Para Dreyfus y Taylor esto es señal de la mecanización de la mente. Las ideas cartesianas o lockeanas son elementos mediadores entre el ser humano y la realidad.

Todo este modelo es para Wittgenstein una imagen o figura: «Una figura nos ha mantenido cautivos» (Wittgenstein, 2009a No 115). Esta figura es más que un concepto, es más que una teorización, es un trasfondo de comprensión que subyace a nuestras teorizaciones y conceptos. Es una «poderosa imagen de “una mente en el mundo”» (Dreyfus & Taylor, 2016, p. 19). Para Dreyfus y Taylor esta imagen tiene rasgos bien definidos: 1) Una estructura mediacional, el conocimiento se produce a través de elementos mediadores, «a través de», no habría un contacto directo con la realidad; 2) en el conocimiento se darían elementos definidos y explícitos últimos (o primeros); 3) no se puede ir más allá de estos elementos; 4) finalmente, detrás de esta imagen hay un contexto dualista, un interior mental y un exterior corporal (Dreyfus & Taylor, 2016, p. 32 ss.) que incluso en el materialismo seguiría considerándose (Cfr. Dreyfus & Taylor, 2016, p. 34). Para Dreyfus y Taylor esta imagen trasciende el racionalismo y el empirismo modernos y persiste incluso en los giros lingüístico, materialista y crítico moderno-contemporáneos (Cfr. Dreyfus & Taylor, 2016, p. 19ss).

Esta idea de una «mente en el mundo» se ha integrado en parte al sentido común. La imagen cartesiana de una mente que se encuentra separada del mundo y de su cuerpo se perpetúa en parte en «el cerebro en una cubeta» de Searle (Cfr. Dreyfus and Taylor, 2016, pp. 23-24) o de forma más masiva por la idea que hizo popular la película Matrix [1999] (Cfr. Dreyfus & Taylor, 2016, p. 161 ss). En ella una super-máquina controla todos los receptores del cerebro, haciendo vivir una realidad ficticia. Dreyfus y Taylor comparan la imagen cartesiana del genio maligno con el «científico maligno», en el ejemplo de Searle (Cfr. Dreyfus & Taylor, 2016, pp. 23-24), y podemos añadirle «la máquina maligna» en el caso de la película Matrix. Mente y realidad se separan radicalmente generando una realidad dual. Su fundamentación filosófica nos sentimos tentados a buscarla más atrás que Descartes, en el platonismo antiguo o en el neoplatonismo medieval, pero en Platón y Agustín el ser humano, su alma, estaba abierta a la verdadera realidad, a aquella que se encontraba más allá de lo sensible, es más bien en el pensamiento moderno donde se opera esta radical separación.

Esta imagen «que nos mantiene cautivos» logra persistir en el esquema computacional de la mente criticado por Dreyfus y Taylor. Los rasgos que veíamos antes se ponen de manifiesto en este modelo: el «a través de», vale decir, el componente mediacional, se expresa en los *inputs* que luego serán procesados, de igual modo es posible reducir la información del computador a elementos claramente definidos: *bits* de información: estos serían elementos últimos. Finalmente, en el computador hay una diferencia entre un elemento claramente material (*hardware*) y otro distinto: las operaciones formales que procesan la información y trabajan con los *inputs*. Esto está a la base de las ideas de Turing, la concepción de que un sistema formal pueda ser ejecutado por una máquina (Cfr. Dreyfus & Taylor, 2016, p. 41). Este modelo puede aplicársele al ser humano y entender que la mente humana funciona de esta manera. Desde una concepción claramente materialista se comprende que la mente ejecuta operaciones formales para procesar la información. Esta última es posible descomponerla hasta llegar a elementos últimos. Esto tiene su correlato en lo que se ha llamado las «Ciencias Cognitivas», especialmente en su versión psicológica.

## 7. Modelos simbólico y neural de la IA

Nos referiremos de manera más específica, aunque todavía muy esquemática, a los dos grandes paradigmas computacionales que han guiado la investigación y el desarrollo tecnológico en el ámbito de la IA. Ya en la década de los cincuenta se perfilan dos modelos que sirven como paradigma del diseño de nuevos computadores: el primero como un sistema para manipular símbolos mentales, el segundo como medio

para modelizar el cerebro (Cfr. Dreyfus & Dreyfus, 1993, p. 25). El primero, llamado "modelo simbólico", tiene para Dreyfus una imagen de ser humano – entendido como ser cognosciente – y en particular de la mente que ha llamado, junto con Taylor, el "modelo computacional de la mente". El segundo, llamado "modelo estadístico neural", tiene una imagen de ser humano que se acerca, para el filósofo norteamericano, aunque todavía de manera muy imperfecta y solo en algún sentido, al modelo comprensión-interpretación que tiene su antecedente fundamentalmente en la filosofía heideggeriana.

La diferencia entre estos dos modelos podemos resumirla en el título del artículo publicado por Hubert y Stuart Dreyfus (1993): «Fabricar una mente versus modelar el cerebro». El primer modelo busca «fabricar una mente a través de estructuras formales que manejen símbolos con el objetivo de resolver problemas. El segundo busca «modelar el cerebro». Uno de sus primeros representantes, Frank Rosenblatt (Cfr. Rosenblatt, 1962), buscó construir un dispositivo físico (Perceptrón), o al menos simularlo, que pudiera generar sus propias habilidades a través del modelamiento de la dinámica de aprendizaje de las neuronas.

### a. El modelo simbólico y su trasfondo filosófico

Con respecto al primer modelo, en 1955, Allen Newell y Herbert Simon tuvieron la idea de que el cerebro humano y la computadora, no obstante las diferencias, podían considerarse dispositivos que generaban conducta inteligente, manipulando símbolos por medio de reglas formales (Cfr. Dreyfus & Dreyfus, 1993, p. 26). Este modelo, para Newell y Simon, remontaría sus fundamentos a teóricos como Frege, Russell y A. N. Whitehead. Los Dreyfus señalan que en realidad esta tradición se remonta más atrás en el tiempo, desde Descartes, Hobbes y Leibniz y podría considerarse que llega hasta el Wittgenstein del *Tractatus* (Cfr. Dreyfus & Dreyfus, 1993, p. 27). Vale decir, una tradición atomista, en tanto considera que el pensamiento humano está basado en elementos claramente definidos y últimos que ingresan a la mente humana para luego ser procesados. Para Dreyfus y Taylor un «modelo mediacional».

Para Descartes la comprensión humana se basaba en representaciones que, a su vez, podían descomponerse en elementos simples. Hobbes, por su parte, pensaba que en la mente se daban componentes formales relacionados por operaciones sintácticas, de esta forma, era posible pensar en un razonamiento semejante al cálculo. Leibniz, por otro lado, tiene la idea de una matices universal que pretendía formalizarlo todo; él pensaba que sería posible asignar signos o números a cada objeto y a los conceptos descomponerlos en elementos simples y últimos (Cfr. Dreyfus, 2007, p. 1137; Cfr. Dreyfus & Dreyfus, 1993, p. 27). El Wittgenstein del *Tractatus*, define el mundo como la totalidad de los hechos atómicos lógicamente independientes (Cfr. Wittgenstein, 2009c 1,1). Tales hechos son combinaciones de elementos primitivos (Cfr. Wittgenstein, 2009c, 2,01 y 2,0124). Así las cosas, desde estos paradigmas, es posible considerar la IA como un intento de encontrar elementos primitivos y vincularlos a través de relaciones lógicas en un dispositivo.

El modelo simbólico, cercano a la tradición filosófica racionalista, en la segunda década del siglo XX dio grandes frutos desplazando en gran medida a su contraparte. Pese a ello, Hubert Dreyfus opinaba que “al combinar el racionalismo, el representacionismo, el conceptualismo, el formalismo y el atomismo lógico en un programa de investigación, los investigadores de la IA habían condonado su empresa a recrear un fracaso” (Dreyfus, 2007, pp. 1137-1138). Tras un largo período el modelo simbólico se impuso, sobre todo después del libro de Marvin Minsky y Seymour Papert, *Perceptrons; an Introduction to Computational Geometry* (Cfr. Minsky & Papert, 1969), en el que se exponía las principales dificultades del modelo estadístico neural. Paradójicamente, por esto último, este libro se considera como un hito importante en el desarrollo de este último modelo. El paradigma estadístico-neural vino a resurgir en la segunda década del siglo XXI y ha tenido uno de sus grandes éxitos en los *LLM* y en la conducción autónoma de vehículos.

Lo que se echa en falta en el modelo simbólico es el trasfondo de sentido o mundo en su manera de concebir aquello que quiere simular: la inteligencia humana. Se subentiende que el ser humano es un ser cognosciente, cuyo modo de captar la realidad es a través de elementos atómicos – Vgr., impresiones sensibles, ideas simples – que luego son procesados. Lo que más arriba señalábamos como tradición racionalista<sup>4</sup> y que se corresponde con el modelo computacional del conocimiento. La crítica de filósofos como Taylor y Dreyfus es que no se toma en cuenta el trasfondo de sentido y su relación fundamental con una praxis humana que tiene elementos opacos, epistemológicamente hablando. Sin embargo, como veíamos, este modelo dio frutos abundantes, sobre todo en los inicios de la IA.

<sup>4</sup> Tomamos este concepto de Charles Taylor (Véase Taylor, 1997, p. 91). Taylor ve en este modelo una consecuencia antropológica que proviene de ontologizar un procedimiento racional de tipo epistemológico. Una determinada concepción de ser humano se consolida: aquella lo considera como un sujeto de conocimiento y no como un ser involucrado en la praxis. Este tipo de racionalismo incluiría tanto el racionalismo cartesiano como el empirismo lockeano, entre otros.

### b. El modelo estadístico-neural

La postura estadístico-neural de la IA «se veía a sí misma como una neurociencia idealizada y holística» (Dreyfus & Dreyfus, 1993, p. 26) y no fue relacionada con ninguna tradición filosófica, sino con la neurociencia. No obstante, Dreyfus ve en este modelo una postura más cercana al paradigma de comprensión-interpretación heideggeriano, aunque aún sujeta a las limitaciones propias de la máquina que no es un estar-en-el-mundo. Los algoritmos estadísticos aprenden de datos proporcionados, simulando de manera imperfecta un razonamiento inductivo (Cfr. Lévy, 2022, p. 12). En sus comienzos se inspiraba en la obra de Donald O. Hebb, considerado un pionero de la psicobiología (Cfr. Hebb, 1949), quien sugirió que un grupo de neuronas podía efectivamente aprender al excitarse simultáneamente a las neuronas A y B, reforzando la conexión entre ambas. Frank Rosenblatt, siguiendo este modelo, pensó la IA como una automatización de los procedimientos neuronales necesarios para discriminar patrones y dar una respuesta coherente (Cfr. Dreyfus & Dreyfus, 1993, pp. 28-29; Cfr. Rosenblatt, 1962).

Pierre Lévy, desde una mirada más actualizada, entiende que «la rama estadístico-neural de la IA implica entrenamiento de algoritmos a partir de acumulaciones masivas de datos para permitir reconocimiento visual, de audio, lingüístico y de otras formas de información» (Lévy, 2022, p. 13). Este modelo, hoy en día, aprende de los datos proporcionados y desemboca en estructuras matemáticas que se adaptan a la realidad manifestada por los datos. En cambio, la formalización lógica del modelo simbólico no aprende de los datos proporcionados, sino más bien depende de un dominio de conocimientos proyectado por ingenieros<sup>5</sup>.

El modelo neural es de carácter mucho más holístico que su contraparte, el modelo simbólico, sobre todo con la gran cantidad de capas que el desarrollo tecnológico de los últimos años ha permitido, ellas se correlacionan entre sí dando lugar a estructuras complejas que interactúan unas con otras. Las relaciones que puede hacer son impresionantes, no obstante, sus dificultades (Cfr. Lévy, 2022, p. 13 ss.). La índole holística se ve reflejada en programas como los *LLM* – que desarrollaremos más adelante – en su capacidad de no solo analizar palabras de manera aislada, sino que logran relacionar palabras, frases, oraciones e incluso estilos de escritura dentro de un contexto determinado. Un cambio en una variable no solo afectaría a un elemento sino al conjunto. Esto les hace generar textos con significados complejos a través de patrones del lenguaje.

El modelo estadístico-neural, por otro lado, tiene una diferencia fundamental con el modelo simbólico, integra datos de entrenamiento: lo que desde el otro modelo se entenderían como *inputs* y *outputs*, resultados de entrada y de salida que el sistema recibe y que se encarga de generar relaciones entre ellos. Como vemos el dato ahora entra en relación y se generan interacciones complejas. Estas se constituyen en el principal aporte del sistema, al menos en su período de entrenamiento.

### c. El problema del sentido común en la reflexión de Dreyfus sobre la IA

Como señalábamos, para Hubert Dreyfus, el modelo estadístico neural se acerca en algunos aspectos al paradigma humano de comprensión-interpretación. No obstante, para él, ambos modelos han tenido como límite el sentido común humano, tan importante para el análisis heideggeriano de la cotidianidad y del uno. El sentido común es demasiado complejo y se resiste a una formalización total (Cfr. Dreyfus, 2007; Cfr. Dreyfus & Dreyfus, 1993). El proyecto de la matices universal de Leibniz ha chocado con un muro. Para Dreyfus el sentido común no es posible descomponerlo en elementos más pequeños, porque – siguiendo a Wittgenstein – este tendría como base una práctica. Como veíamos, autores como el segundo Wittgenstein y Heidegger tienen una concepción holística de la comprensión humana que considera el entorno y las relaciones entre los distintos elementos, que se distancia de la mirada racionalista. Considerar que es posible descomponer la comprensión en elementos últimos parece un despropósito (Cfr. Wittgenstein, 2009a, pp. 206-207. N° 46-47). En el caso de Heidegger, no es posible descomponer la comprensión de la cotidianidad en elementos últimos. Todo depende del contexto de significatividad que otorga el mundo.

Pese a ello, Edmund Husserl, maestro de Heidegger, no renuncia a la idea de que el mundo – tal como él lo entendía – se torne explícito, incluso dejando abierta la posibilidad de formalizar algo como el sentido común. No obstante, Husserl se enfrentó con la complejidad de esta empresa, sobre todo tras las críticas de Heidegger (Cfr. Dreyfus & Dreyfus, 1993, p. 41). Recordemos que para Husserl la conciencia posee intencionalidad, vale decir, la conciencia es siempre conciencia de algo. En su estructura distingue entre *noesis* y *noema*, la primera es la vivencia y la segunda es el contenido de esa vivencia, la intencionalidad es entonces la correlación entre *noema* y *noesis*. Por ejemplo, se ama (*noesis*) lo amado (*noema*), se representa lo representado, etcétera. El acto intencional, para Husserl, posee esa direccionalidad que es la intencionalidad en la medida que el *noema* posee significado, este significado implica predelineamientos que describen los rasgos estables, que permanecen “inviolablemente los mismos”, así como otros de

<sup>5</sup> Lévy ve como insuficientes ambos modelos y propone complementarlos a través de un metalenguaje llamado *Information Economy Meta Language* (IEML) (Cfr. Lévy, 2022, p. 19 ss.)

rasgos posibles, pero no estables. Esto ocurriría dentro de un contexto que determinaba la inteligencia en tanto proveía un horizonte de expectativas o predelineamientos. En definitiva, un mundo y sus relaciones constitutivas. Para Hubert Dreyfus hay en la reflexión de Husserl una semejanza con el intento de la formalización de todo por parte de la IA, al punto de llamarlo "abuelo de la IA" (Cfr. Dreyfus & Dreyfus, 1993, p. 39), en tanto se le considera la culminación de la tradición cartesiana. Para Dreyfus en cambio, no es posible formalizar todo, en la base de nuestro pensar hay prácticas, no elementos definidos

Algo similar ve Dreyfus en Marvin Minsky. En el contexto de la IA, él buscaba generar un marco (*frame*), que iría muy en la línea de proporcionar un contexto, algo así como un tipo de mundo previo, entendido como una estructura de datos que representa una situación estereotipada que sirve de trasfondo, como "la de estar en cierta clase de habitación, o la de ir a la fiesta de cumpleaños de un niño" (Minsky citado desde Dreyfus & Dreyfus, 1993, p. 40; Cfr. Minsky, 1981). Esto podría pensarse que va en la línea de un trasfondo, mundo o contexto. Este marco es una red de nodos y relaciones que posee niveles superiores e inferiores. Los superiores serían equivalentes a lo que Husserl entendía como rasgos estables y los inferiores corresponderían a asignaciones por defecto, rasgos esperables en situaciones normales. Pese a que toma en cuenta un contexto o trasfondo, lo que Dreyfus critica a esta postura es que considera que es posible resolver de esta forma el problema del sentido común. Un punto que para Dreyfus es complejo es el problema de la relevancia, ¿cómo sabría la computadora cuáles hechos eran más relevantes que otros en una situación dada? Lo que Dreyfus echa en falta es la plasticidad del mundo. Para Minsky esto sería posible resolverlo a través del manejo de unos pocos millones de datos sobre los objetos y sus funciones (Dreyfus, 2007, p. 1138). Dreyfus se pregunta: ¿cómo determinaría el programa que algo cambió en el mundo? En una situación dada tendría que haber una batería de marcos intercambiables, teniendo que la computadora seleccionar el cuadro más relevante. Minsky a comienzos de este siglo reconoce el límite que tiene la IA con respecto al sentido común y el estancamiento que venía ocurriendo desde los años setenta (Dreyfus, 2007, p. 1139).

En un artículo publicado el 2007 –«Failed and How Fixing It Would Require Making It More Heideggerian» ("Por qué falló la IA heideggeriana y cómo solucionarla requeriría hacerla más heideggeriana") – Hubert Dreyfus entiende que la IA sería más exitosa si acentuara su impronta heideggeriana. Para él, el gran problema era modelar el sentido común y, por tanto, un trasfondo o mundo en sentido heideggeriano que fuera más plástico. Dreyfus nos relata algunos intentos entre los años sesenta y setenta en que los programadores se abocaron a "micromundos". Esto es todavía insuficiente para Dreyfus, pues no apuntan al mundo real, sino a pequeños contextos. Más adelante, señala Dreyfus, la crítica de Heidegger es implícitamente aceptada, y destaca tres intentos de modelar una "IA heideggeriana": El enfoque conductista de Rodney Brooks (Cfr. Brooks, 1986), el modelo pragmático de Phil Agre (Cfr. Agre, 1997) y el modelo neurodinámico de Walter Freeman (Cfr. Freeman, 1991). En las dos primeras ve Dreyfus el problema que no se genera ni aprendizaje, ni habilidades: el robot (*animat*) de Brooks solo responde a características fijas y aislables del entorno que capta a través de sensores (Cfr. Dreyfus, 2007, p. 1140); el programa Pingo – de Agre y David Chapman, que simula un agente que interactúa con el juego Pingo – trata de reaccionar a posibilidades de acción y no a objetos, pero el problema, para Dreyfus, es que objetiva las funciones de los objetos y las posibilidades de acción (Cfr. Dreyfus, 2007, pp. 1142-1144).

Más allá de estas críticas, Dreyfus destaca el trabajo neurocientífico de Walter Freeman (Cfr. Dreyfus, 2007, p. 1150 ss). Al igual que Merleau-Ponty, Freeman considera que el cerebro se encuentra encarnado en un animal que busca satisfacer sus necesidades en su propio ambiente. En él, el animal no capta características sin significado, que luego procesaría para lograr una representación. El entorno ya es significativo para el animal. Pero ¿cómo seleccionar desde el trasfondo aquello que es significativo? Para ello, las interacciones previas con el mundo son fundamentales para formar patrones: «el sistema de percepción del animal está preparado por la experiencia pasada y la excitación para buscar y ser recompensado por experiencias relevantes». En el caso de un conejo, cuando busca una zanahoria que comer y lo consigue, las conexiones entre neuronas se vuelven más fuertes a través de la sinapsis, así los conjuntos de células no son simplemente receptores pasivos, sino que están sintonizados para responder a lo que es significativo. Otro ejemplo, las células del conejo involucradas en el escape de un zorro se conectan entre sí generando un conjunto unido de células, así las neuronas del conjunto y el bulbo olfatorio están preparados para una respuesta específica cuando siente el olor de un zorro. No es que el órgano olfativo responda a ciertas características, no se procesan estímulos que luego se convierten en representaciones, a partir de las cuales el cerebro inferiría lo que se encuentra en el entorno. Más bien, se forman «estados energéticos» que están abiertos a posibles atractores. El cerebro animal forma cuencas de atracción y el conjunto de estas cuencas se denomina «paisaje atractor» (*attractor landscape*). Una analogía es una nave volando sobre un valle lunar lleno de cráteres, las zonas bajas de estos últimos serían los atractores, la nave espacial va buscando un fondo de cráter dónde aterrizar. Al respecto, se le presentan atractores diferentes para cada clase de estímulo. Pero los patrones aprendidos no se relacionan directamente con el estímulo, sino con su importancia, en este caso el mejor sitio para aterrizar. Un estímulo nuevo, anteriormente no significativo, establece un nuevo atractor y ordena las demás cuencas del paisaje

atractor. Cuando, por ejemplo, el animal aprende un nuevo olor, hay un cambio en todos los demás patrones, no se opera con representaciones fijas, sino con significados. En este sentido existe un arco intencional y, de acuerdo con él, las experiencias en el mundo no son idénticas. El animal toma en cuenta cómo van las cosas y busca continuar con un camino prometedor, y si la acción no va por buen camino, el cerebro se auto organiza para saltar a otro atractor. En el caso del olfato, funciona con otros sistemas cerebrales. No es que el bulbo olfatorio envíe un mensaje a la parte del cerebro apropiada. Tal como Merleau-Ponty, Freeman piensa que lo que nos mueve a actuar es un desequilibrio entre el yo y el mundo. En esto hay que tener en cuenta que lo que representa este tipo de modelo no son características del mundo cotidiano, sino más bien representaciones de estado cerebrales. Otra cosa para tener en cuenta es que las acciones fisicoquímicas no se desarrollan como secuencias separadas y paralelamente, sino que se conectan a través de vórtices o nudos<sup>6</sup>. Una gran cantidad de experiencias de estímulos y respuesta se interconectan involucrando a todo el sistema. En cuanto a su aplicación a la IA, Freeman señala que los cerebros no usan números como símbolos, pero sí usan eventos discretos en el tiempo y en el espacio, que podrían representarse mediante números. Dreyfus destaca que Freeman ha utilizado su modelo neurodinámico para simular la actividad del cerebro de una salamandra, específicamente modelando sus necesidades de autoconservación y búsqueda de alimento.

Para Dreyfus, las otras propuestas «heideggerianas», arriba mencionadas, no resuelven el problema del marco – o contexto de sentido –, sobre todo en cuanto al problema de la relevancia y el cambio en las condiciones de este. Vale decir, las condiciones de un mundo o trasfondo, y la relevancia de las situaciones en él. En todo caso, el logro de Freeman, para Dreyfus, se encuentra lejos del modelamiento del afrontamiento humano de la realidad. Consideramos que, si bien aún existe lejanía con el afrontamiento humano, la conducción autónoma de vehículos ha logrado superar parcialmente algunas de las dificultades que planteaba Dreyfus, sobre todo el aspecto de la relevancia, esto a través del manejo estadístico.

Algo importantísimo que señala Dreyfus del trabajo de Freeman es la relación que se establece entre actividad neuronal, es decir la excitación neuronal que genera una cadena de conexiones sinápticas, y la explicación fenomenológica de la percepción y acción cotidianas: Freeman «ha mostrado lo que hace el cerebro para proporcionar el substrato material para la explicación fenomenológica de la percepción y la acción cotidiana de Heidegger y Merleau-Ponty»<sup>7</sup> (Dreyfus, 2007, p. 1157). En particular la respuesta aprendida ante un estímulo en el marco de un contexto o entorno. La respuesta del animal se va configurando en hábitos y esto tiene su correlato en las conexiones neuronales que se forman. De esta manera, Dreyfus ve con buenos ojos el modelo neuronal de la IA, este intentaría traducir la actividad humana (y animal) a un tipo de conexiones que simulan las conexiones neuronales, base física de la acción.

Con ello, para Dreyfus, se ha dado un avance – que si bien no viene de la filosofía, sino de la neurociencia – respecto de un paradigma que considera la conducta y el aprendizaje humano desde el esquema comprensión-interpretación.

## 8. Los LLM y la interacción con el ser humano

La comprensión humana, no es algo posible aún para la IA. No obstante, al interactuar con los *LLM* ("Large Language Model") – del que el chat GPT, en sus diferentes versiones es uno de sus representantes más famosos<sup>8</sup> – nos da la impresión que la IA se ha acercado mucho. La máquina responde a nuestras interrogantes con coherencia semántica y con textos, que, aunque no siempre exentos de error, son relativamente acertados. En el interesantísimo artículo – "Large Language Models, Meaning, and the Patterns of Human Language Use" (Durt & Fuchs, 2024) [“Grandes modelos lingüísticos, significado y patrones de uso del lenguaje humano”] – de Christoph Durt y Thomas Fuchs, se argumenta que no es el modelado de la actividad mental lo que se persigue en los *LLM*, sino más bien simular los patrones comunes del uso del lenguaje. Esto es importante, pese a seguir el modelo neuronal de la IA, no se persigue simular la actividad cerebral, sino manejar las regularidades del lenguaje, considerándolo un sistema que se da dentro de un contexto. La peculiar comprensión humana se deja de lado y se intenta trabajar con un producto humano, los textos de la red o aquellos textos hablados que pueden ser traducidos al lenguaje escrito. El lenguaje tiene, como veremos, independencia del individuo que interpreta dichos textos. Los *LLM* tratan de manipular el lenguaje humano a través de patrones estadísticos. Detrás de la elaboración de textos se encuentran patrones que ordenan y predicen el lenguaje escrito y de esta manera es posible generar textos coherentes y comprensibles.

<sup>6</sup> Afirmó Merleau-Ponty: “Sólo hay que admitir que las acciones físico-químicas de que el organismo está en cierta manera compuesto, en lugar de desarrollarse en secuencias paralelas e independientes, como lo querría el espíritu anatómico; en lugar de entremezclarse en un conjunto donde todo dependiera de todo y donde ningún clivaje fuera posible, se constituyen, según la expresión de Hegel, en ‘nudos’ o en ‘torbellinos’ relativamente estables – las funciones, las estructuras del comportamiento –, de tal manera que el mecanismo se dobla en una dialéctica.” (Merleau-Ponty & Waelhens, 1957, p. 217)

<sup>7</sup> Las cursivas son nuestras.

<sup>8</sup> También es conveniente considerar programas como Gemini, Llama y Claude, entre otros.

Como vimos, para Hubert Dreyfus estamos lejos de llegar a simular la comprensión humana tal como era entendida por Heidegger y en particular del sentido común, la que está relacionada con el fenómeno del uno, desde donde parte toda comprensión e interpretación. Para el filósofo norteamericano, el problema de fondo sería saber qué hechos son relevantes en una situación determinada, partiendo de un marco (*frame*) o contexto y, sobre todo, la imposibilidad de la total formalización del sentido común. Pero podemos preguntarnos si es posible saltarse la necesidad de imitar la comprensión humana. ¿Habrá otro camino para llegar a formular textos coherentes semánticamente? ¿Podría la máquina captar la relevancia de una situación o de un texto sin necesidad de comprender? La IA, y en general la técnica, nos han sorprendido con un gran despliegue de creatividad. Una de las alternativas que se tiene para poder superar estos problemas es cambiar el enfoque. En el caso de los *LLM* no se busca comprender y expresarse lingüísticamente como lo hacen los seres humanos, sino manejar los patrones del lenguaje escrito (o del lenguaje hablado transformado en escrito). En la red se tiene un impresionante acervo de textos construidos por el ser humano desde hace décadas. Lo que han hecho los *LLM* es manejar estos a través de patrones de lenguaje, sin necesidad de tener comprensión al modo humano, y lograr establecer regularidades entre ellos. Así, y bajo los patrones ortográficos, sintácticos y semánticos, les fue posible construir respuestas coherentes. Palabras, expresiones e ilaciones se han transformado en números y se ha logrado manejarlos de manera estadística. Dicho de manera muy tosca: los *LLM* han logrado ordenar lo que nosotros mismos hemos escrito.

Si el acento de los *LLM* está más bien en el trabajo con los patrones del lenguaje y no con el modelamiento de la mente, entonces, consideramos, y he aquí nuestro aporte a la reflexión, que es posible, a través del lenguaje, llegar al contexto y con ello a cierto sentido común – aun cuando este último sigue siendo indescifrable en su totalidad –. Esto porque no es posible hacer una separación tan radical entre lenguaje (en este caso escrito) y significado, considerando al primero como un mero sistema formal. Si pensamos, acudiendo al Heidegger posterior a *Ser y tiempo*, que el lenguaje no depende solo del *Dasein* individual, de ahí aquella oscura expresión, “el lenguaje habla” (Cfr. Heidegger, 1990, p. 18 y 30), podemos pensar en la posibilidad de una relativa independencia del lenguaje respecto al sujeto que conserva estructuras y patrones propias de un lenguaje humano y su significatividad. Para el filósofo alemán, el lenguaje trasciende al *Dasein* individual y tiene independencia respecto de él. Podemos decir que este posee una “existencia” o “vida” propia que regula el hablar del *Dasein* individual. Con esto Heidegger pretende superar la concepción del lenguaje como un fenómeno subjetivo. Incluso podemos interpretar que el lenguaje trasciende a las comunidades particulares en que nos desenvolvemos y regula su comprensión. Al separar al lenguaje de su uso podemos conservar en parte su significatividad, para insertarlo nuevamente en su uso. En el paréntesis que ocurre en la separación de su uso, vale decir, en el trabajo que hacen de él los *LLM*; los textos se transformarían en números y se manejarían con patrones estadísticos.

Por lo mismo, el trabajo de los *LLM* no es suficiente para lograr el objetivo para el que fueron creados, se hace necesaria la interpretación humana para darle sentido a los textos producidos y hacer que efectivamente logren ser comprendidos. Aunque los *LLM* generen textos coherentes y relativamente correctos, si no hubiera seres humanos que los interpretaran y les dieran el sentido, los textos no cumplirían la función para la que fueron creados, vale decir, ser interpretados por el ser humano. La máquina trabaja con números y estadística, pero el sentido de los textos, su comprensión e interpretación, la aportamos nosotros. Es el ser humano que contribuye con su comprensión sobre la base de textos cada vez más coherentes y bien construidos. Por ejemplo, un texto de tipo periodístico, digamos sobre una noticia específica, es generada por los *LLM* recurriendo a regularidades que encuentra en la red, sin comprender su sentido. Su interpretación y sentido se la damos nosotros como seres humanos que estamos inmersos en un contexto y, en un sentido mucho más amplio, inmersos en el lenguaje. Por tanto, no es necesario que la máquina, al menos en lo que respecta a los *LLM*, comprenda como lo hace un ser humano, para generar textos que a nosotros nos parecen coherentes. El que puedan ser desconectados, reconfigurados y reconectados al uso muestra la independencia del sujeto respecto de la conservación de la significatividad, pero a la vez una dependencia en cuanto necesita de un lector para ser interpretado y trasmitir un sentido. Por un lado, el lenguaje trasciende el uso individual y regula nuestra comprensión, pero por otro necesita de la interpretación humana para que cumpla su función.

Al desviar el enfoque del cerebro-mente al lenguaje, en el sentido que se deja de buscar que la máquina comprenda y se persigue más bien manejar patrones de lenguaje y regularidades en los textos, consideramos que se está realizando un acierto. El lenguaje escrito, vale decir textos de la red, “separado de su uso” en cuanto solo es manipulado por la máquina, puede trabajarse para volver a insertarlo en su uso en tanto interpretado por seres humanos. El lenguaje por tener esa independencia del *Dasein*, conserva algo de su contexto en los textos de la red, el cual puede ser captado nuevamente por el lector. Ciertamente, la comprensión última está en el *Dasein* y en su comunidad, pero esto no quiere decir que el manejo estadístico de un lenguaje, que procede del uso humano, inserto en su mundo o contexto, no logre preservar algo de él y que este no pueda complementarse con los mundos en los que está inserto el que interpreta estos textos. Si pensamos en contextos como el periodístico o el científico, estamos pensando

en mundos concretos, cuyo sentido y significado, hicieron posibles textos que nos hemos podido trasmisir desde que estas actividades existen. ¿Por qué pensar que la significatividad no esté impregnada en algo como el lenguaje escrito, trascendiendo nuestra mente o nuestro encarnado e involucrado estar-en-el-mundo? Ciertamente no todo el significado del lenguaje es posible capturarlo con medios estadísticos (Cfr. Durt & Fuchs, 2024, p. 116). Pero si la IA sabe aislar ciertos contextos o temas, puede perfectamente moverse en ellos y generar estadísticamente respuestas coherentes. Ciertamente la complejidad de un mundo que no es solo lenguaje escrito y de un lenguaje que está basado en prácticas, no queda en entredicho. Algo como la comprensión humana todavía no es posible a través de la IA, pero esto no quiere decir que no haya patrones en nuestras expresiones lingüísticas que estén transidos de mundo. Evidentemente, la IA aún no es capaz de captar la riqueza del mundo y toda la complejidad del sentido común.

Es necesario señalar que el modelo estadístico-neural ha ido desarrollando, con mucha eficacia, algo que Dreyfus echaba en falta en los primeros intentos de generar lo que él llamó una IA heideggeriana – vale decir, una IA que integrara el mundo, o al menos el o los contextos, a la manera como el *Dasein* comprende las cosas desde el mundo –, nos referimos al aprendizaje. Fundamentalmente, a través del modelamiento estadístico, los nuevos programas logran encontrar patrones de manera inductiva que le permiten clasificar y manipular la gran cantidad de datos que manejan.

## 9. Conclusión

La comprensión humana está transida de significatividad, incluso la experiencia más básica, la percepción, no está exenta de ella (Merleau-Ponty). Ello supone un trasfondo o mundo, que consiste en un plexo de relaciones compartido que da significatividad a los entes que en primera instancia son captados como útiles (Heidegger). Esto es lo que echa en falta Hubert Dreyfus en el modelo simbólico de la IA, que se basa para él y Taylor en “el modelo computacional de la mente”. Vale decir, en un modelo que cosifica sus entradas (ideas simples, impresiones sensibles o similar) y cuyas relaciones son establecidas por el sujeto cognosciente. En esto estriba el acierto del modo de acceso heideggeriano al estudio del *Dasein*, que considera a este fundamentalmente un sujeto absorto en la praxis de la cotidianidad y en un segundo momento un sujeto cognosciente, al contrario de la tradición filosófica que piensa a este último paradigma como el central. El paradigma cognoscitivo con su esquema sujeto-objeto entiende la realidad como objetos separados que luego entran en relación, al contrario del modelo heideggeriano que considera que la comparecencia del ente se da gracias a las relaciones o remisiones constitutivas del mundo. Ellas constituyen su significatividad. De manera más originaria el *Dasein* capta los entes como útiles – es decir en su interrelación – y en un segundo momento como objetos de conocimiento que son entes descontextualizados. Por todo esto, el esquema heideggeriano marca un modelo mucho más holístico de comprensión. Es su carácter holístico-relacional lo que Dreyfus consideraba prometedor del modelo estadístico-neural, pese a que no ha resuelto del todo el problema de la comprensión, ni se acerca suficientemente a toda la complejidad del mundo y del sentido común. Este último remite al existencial del uno heideggeriano, que además de tener un carácter inauténtico, posee una índole opaca, epistemológicamente hablando, porque se basa en última instancia en prácticas y no en pensamientos. Como veíamos es en este existencial que Dreyfus ve la trasmisión del trasfondo de significatividad, aquel que es la base del entendimiento cotidiano.

Por otro lado, detrás de todo nuestro pensamiento y lenguaje no se encuentran elementos atómicos, sino formas de actuar (Wittgenstein, Heidegger), vale decir prácticas. Estas últimas poseen significatividad, tanto es así, que es a través de estas prácticas que se nos trasmite todo un trasfondo de sentido que nos posibilita comprender la realidad (Heidegger, Dreyfus y Taylor). Si la inteligencia artificial quiere, de algún modo, emular la comprensión humana debe hacerse cargo de un esquema de comprensión más complejo que el que ha presentado la tradición racionalista y empirista – o los epígonos contemporáneos que critican Dreyfus y Taylor – al menos así lo ha entendido Hubert Dreyfus. Este último, nos relató algunos intentos de hacer una IA más holística – él la señaló como “heideggeriana” – que tuvieron cierto éxito relativo (Dreyfus, 2007).

Como hemos visto, el modelo estadístico-neural incluye el aprendizaje de la máquina, a través de datos de entrenamiento con los que va modelando su acción. Esto muestra una índole adaptativa que es propia de su paradigma y que supone un adecuamiento a un contexto. Esto es para Dreyfus un camino que se ve auspicioso, si bien aún no del todos satisfactorio. El hecho de considerar un contexto, que dé significatividad y que considere la realidad en su interrelación constitutiva, es un avance respecto del modelo simbólico. Este avance implica establecer relaciones a través de los datos de entrenamiento de la máquina. Lo que antes se conocían como entradas y salidas (*inputs* y *outputs*), son ahora los datos básicos que la máquina toma para establecer relaciones entre ellos. Lo que se busca, al menos en el nivel de entrenamiento, son las relaciones que los constituyen como datos.

Efectivamente, hay cosas que la máquina no puede hacer. Una de ellas es la real comprensión, a la manera humana, que supone un trasfondo de sentido ligado a prácticas. Ciertamente se han dado avances, en tanto que el enfoque estadístico neural es mucho más holístico. Este ha tratado de simular la dinámica de aprendizaje neuronal, que, para Dreyfus, sería la base material de nuestro encarnado estar-en-el-mundo.

Más allá de ello, el modelo de los *LLM*, siguiendo dentro de la dinámica holística neural, tiene un enfoque un tanto distinto. No es la comprensión humana la que trata de emular, sino que persigue trabajar con el lenguaje, manejando sus patrones. Gracias a la inmensa capacidad que han logrado los ordenadores, a estos programas les es posible captar los patrones que se encuentran en los escritos de la red e identificar estos en sus diferentes contextos. Siendo semejantes a los predictores de texto, pero mucho más sofisticados, logran elaborar escritos coherentes y, en muchas ocasiones, acertados. Consideraremos que en estas elaboraciones se conserva parte de un trasfondo de sentido, al menos el implícito en un texto con un tema determinado. De hecho, todos los textos que leemos y elaboramos necesitan de un trasfondo para ser interpretados y este, en parte, se conserva tras la intervención de los *LLM*. Un texto coherente conserva algo de este trasfondo de sentido, a pesar de que el lenguaje sale de su uso para ser manejado estadísticamente. Si, como nos decía Heidegger, “el lenguaje habla”, este no depende solo del sujeto que lo habla, lee o escribe; sino que tiene cierta independencia frente a él. Al sacar los *LLM* el lenguaje de su uso para ser codificado y manejado estadísticamente, lo hace con el objeto de volver a insertarlo en su uso. Es evidente que todo esto tiene sentido en la interacción con los humanos. Sin estos los textos no serían comprendidos, porque los *LLM* no tienen esa capacidad. Los *LLM* manejan un lenguaje que procede del uso humano, por ejemplo, de la práctica del periodismo. Ellos trabajan este lenguaje de manera estadística, convirtiendo sus regularidades en ecuaciones que manejan números. De esta manera ordenan fragmentos de textos y los expresan siguiendo reglas gramaticales y semánticas. Pero estos textos se integran nuevamente al uso humano de la lectura y es ahí donde la significatividad cobra realidad en la interpretación. Es el ser humano desde su mundo o contexto el que interpreta los textos emanados de los *LLM*.

Si comprendemos la IA como «la rama específica de la ciencia de la computación que se ocupa del modelado y la simulación de la inteligencia humana» y su objetivo como aumentar el conocimiento que disponemos (Cfr. Lévy, 2022, p. 13); entonces podemos concluir que no es indispensable que la IA llegue a desarrollar una comprensión como la humana, sino que la simule en la medida de sus posibilidades. Ella ya puede simular razonamientos inductivos y deductivos y actualmente, con las redes neurales es capaz de un tipo de aprendizaje autónomo. Para generar textos que nos sean significativos, no es necesario lograr una conciencia como la humana que posee una comprensión basada en un mundo. Basta con que esta comprensión siga siendo nuestra prerrogativa y que cierta parte del contexto pueda transmitirse a través de textos elaborados sobre la base de nuestras propias reflexiones. La opción de la simulación a través de la manipulación del lenguaje, consideramos que es un camino creativo y relativamente eficiente que no sólo genera textos coherentes, sino que incluso maneja cierto contexto a través de la clasificación de los textos. Cierta significatividad queda inserta en estos últimos, dada la independencia del lenguaje respecto del individuo que luego lee e interpreta. Luego los textos al volver al uso humano pueden ser reinterpretados por un individuo que es un estar-en-el-mundo. Lo que tiene de distinto el enfoque de los *LLM* es que no se centra en la actividad del *Dasein* individual, sino en algo que lo trasciende: el lenguaje humano.

En este trabajo no hemos querido en modo alguno ser exhaustivos. Hay aspectos fundamentales que no hemos considerado, tanto de la comprensión humana como de la IA, entre los primeros algunos existenciales heideggerianos que son fundamentales en la consideración de la comprensión humana, por ejemplo: el cuidado, el ser-para-la-muerte, la disposición afectiva, la temporalidad humana, entre otros. Respecto de la IA hemos esbozado una concepción más bien general, con el propósito de compararla con la comprensión humana. Ciertamente, el actual enfoque sectorial, que se especializa en tareas específicas, no busca una comprensión de este tipo. En este trabajo, solo hemos querido presentar algunos paradigmas ontológicos, epistemológicos y computacionales. Más allá de la conciencia que aún es una meta que parece lejana, la idea de un mundo a la manera heideggeriana es algo aún no alcanzado por la IA. Se han dado avances en cuanto a las relaciones holísticas involucradas que persiguen simular la comprensión humana, sobre todo desde el actual modelo neural, pero todo esto desde una máquina que recibe estímulos y datos y no de un estar-en-el-mundo al modo humano.

Abriendo la discusión más allá de lo hasta aquí reflexionado, y dado que hemos presentado una diferencia entre la comprensión humana y la simulación de la IA, vislumbramos un riesgo en la consideración de esta temática, en el trasplantar, sin mayor previsión, un modelo computacional a nuestra concepción epistemológica y antropológica, sin advertir las diferencias que subsisten entre máquina y ser humano. Caer en ello es una tentación no menor, sobre todo al ver los impresionantes avances de la IA. El sentido común que ha integrado la idea de que somos “como un computador” es un presupuesto que no es fácil

de vencer, pero si nos rendimos a él corremos el riesgo de entender nuestra comprensión humana de manera artificial.

## 10. Referencias bibliográficas

- Acevedo, J. (1999). Acerca del “problema” del conocimiento (Heidegger, Hartmann y Hessen). En *Heidegger y la época técnica* (pp. 21-42). Editorial Universitaria.
- Agre, P. E. (1997). *Computation and Human Experience*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511571169>
- Brooks, R. (1986). A robust layered control system for a mobile robot. *IEEE Journal on Robotics and Automation*, 2(1), 14-23. IEEE Journal on Robotics and Automation. <https://doi.org/10.1109/JRA.1986.1087032>
- Dreyfus, H. L. (1996). *Ser-en-el-Mundo. Comentario a la División I de Ser y tiempo de Martin Heidegger* (F. Huneeus & H. Orrego, Trad.). Cuatro Vientos.
- Dreyfus, H. L. (2007). Why Heideggerian AI failed and how fixing it would require making it more Heideggerian. *Artificial Intelligence*, 171(18), 1137-1160. <https://doi.org/10.1016/j.artint.2007.10.012>
- Dreyfus, H. L. (2017). On the Ordering of Things. Being and Power in Heidegger y Foucault. En M. Wrathall (Ed.), *Background Practices, Essays on the Understanding of Being*. Oxford University Press. <https://play.google.com/books/reader?id=qjooDwAAQBAJ&pg=GBS.PT2>
- Dreyfus, H. L., & Dreyfus, S. E. (1993). Fabricar una mente versus modelar el cerebro: La inteligencia artificial se divide de nuevo (C. Reynoso, Trad.). En S. R. Graubard (Ed.), *El nuevo debate sobre la inteligencia artificial: Sistemas simbólicos y redes neuronales* (pp. 25-58). Gedisa.
- Dreyfus, H. L., & Taylor, C. (2016). *Recuperar el realismo* (J. Carabante, Trad.). Ediciones Rialp, S.A.
- Durt, C., & Fuchs, T. (2024). Large Language Models and the Patterns of Human Language Use. En M. Cavallaro & N. De Warren (Eds.), *Phenomenologies of the Digital Age* (pp. 106-121). Routledge. <https://doi.org/10.4324/9781003312284-7>
- Freeman, W. J. (1991). Fisiología de la percepción. *Investigación y ciencia*, 175, 30-38.
- Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. Wiley.
- Heidegger, M. (1927). *Sein und Zeit*. M. Niemeyer.
- Heidegger, M. (1990). *De camino al habla* (Y. Zimmermann, Trad.). Ediciones del Serbal.
- Heidegger, M. (1997). *Introducción a la metafísica* (Á. Ackermann Pilári, Trad.). Gedisa.
- Heidegger, M. (1998). *Ser y Tiempo* (J. E. Rivera, Trad.). Universitaria.
- Heidegger, M. (2006). *Prolegómenos para una historia del Concepto de Tiempo* (J. Aspiunsa, Trad.). Alianza Editorial.
- Hessen, J. (2006). *Teoría del conocimiento* (J. Gaos, Trad.). Losada.
- Lévy, P. (2022). IEML: Rumo a uma mudança de paradigma na Inteligência Artificial. *MATRIZes*, 16(1), Article 1. <https://doi.org/10.11606/issn.1982-8160.v16i1p11-34>
- Llamas, E. (2001). *Charles Taylor: Una antropología de la identidad*. Ediciones Universidad de Navarra.
- Locke, J. (1999). *Ensayo sobre el entendimiento humano* (E. O'Gorman, Trad.). Fondo de Cultura Económica.
- Marrades, J. (2014). Sobre la noción de “forma de vida” en Wittgenstein. *Agora: papeles de Filosofía*, 33(1). <https://doi.org/10.15304/ag.33.1.1873>
- Merleau-Ponty, M. (1993). *Fenomenología de la percepción* (J. Cabanes, Trad.). Planeta-Agostini.
- Merleau-Ponty, M., & Waelhens, A. (1957). *La estructura del comportamiento* (E. Alonso, Trad.). Hachette.
- Minsky, M. (1981). A Framework for Representing Knowledge. En Haugeland (Ed.), *Mind Design* (MIT Press).
- Minsky, M., & Papert, S. (1969). *Perceptrons—An Introduction to Computational Geometry*. M.I.T. Press.
- Rosenblatt, F. (1962). *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan Books.
- Taylor, C. (1989). Embodied Agency. En H. Pietersma, *Merleau-Ponty: Critical Essays* (University Press of America, pp. 1-21).
- Taylor, C. (1997). *Argumentos filosóficos: Ensayos sobre el conocimiento, el lenguaje y la modernidad* (F. Birués Beltrán, Trad.). Paidos.
- Vattimo, G. (2006). *Introducción a Heidegger* (A. Báez, Trad.). Gedisa.
- Wittgenstein, L. (2009a). Investigaciones Filosóficas (A. García Suárez & C. Moulines, Trad.). En *Obras Wittgenstein I*. Gredos.
- Wittgenstein, L. (2009b). Sobre la certeza (J. Lluís Prades & V. Raga, Trad.). En *Obras Wittgenstein I*. Gredos.
- Wittgenstein, L. (2009c). Tractatus logico-philosophicus (I. Reguera, Trad.). En *Obras Wittgenstein I*. Gredos.