

# Mente natural, mente artificial

Filosofía de las ciencias  
cognitivas

Asier Arias Domínguez



bauplan

# Índice

PREFACIO	11
Capítulo 1. La revolución cognitiva	15
1. Introducción	15
2. La revolución cognitiva	19
3. La IA en el marco de la revolución cognitiva	31
4. La lingüística en el marco de la revolución cognitiva	38
5. Las neurociencias en el marco de la revolución cognitiva	46
Capítulo 2. Los paradigmas históricos	57
1. Introducción	57
2. El simbolismo clásico	59
2.1. La hipótesis del sistema de símbolos físicos	62
2.2. El lenguaje del pensamiento	68
3. El paradigma conexionista	73
3.1. El neuroconexionismo	79
4. La objeción de la sistematicidad	86
5. Propuestas híbridas	88
6. Aprendizaje profundo	91
7. Procesamiento predictivo	95

Capítulo 3. Objeciones filosóficas al computacionalismo	99
1. Introducción	99
2. Las objeciones fenomenológicas de Hubert Dreyfus	100
3. La objeción semántica de John Searle	107
4. La objeción gödeliana de Lucas-Penrose	120
5. La objeción metodológica de Noam Chomsky	125
Capítulo 4. Alternativas teóricas al computacionalismo	131
1. Introducción: el entramado postcognitivista	131
2. Dinamicismo y enactivismo	136
2.1. Dinamicismo	136
2.2. Enactivismo	141
Capítulo 5. Para una filosofía de las ciencias cognitivas	151
1. Introducción	151
2. La batalla interparadigmática	161
EPÍLOGO	171
Referencias bibliográficas	173

*Para Diana*

# Prefacio

Siendo la filosofía, como es, la más feliz y afortunada de las disciplinas, nadie le reprocha que se acerque a sus objetos desde el ángulo que le plazca. Así las cosas, el título de este libro podría suscitar la expectativa de emocionantes exploraciones de los temblores más íntimos de la conciencia de «superinteligencias artificiales» o de los impresionantes poderes cognitivos de las «mentes postbiológicas» de avanzadísimas civilizaciones alienígenas (Schneider, 2019). Dejamos, sin embargo, esas cumbres del arrebató intelectual sin dique ni límite a sus dominios propios: la filosofía de la mente, acaso; la ciencia ficción, con pulso más firme. Aquí concentraremos nuestros esfuerzos en una empresa más modesta: la de esbozar los rudimentos básicos de la filosofía de las ciencias cognitivas.

Dibujar los contornos de esos rudimentos es un trabajo que puede hacerse de muchas maneras, pero en ningún caso llevará muy lejos sin una cierta comprensión previa de lo que son las propias ciencias cognitivas. Así pues, atenderemos a su génesis histórica y a la naturaleza de las principales tradiciones de investigación en el área antes de abordar tareas propiamente filosóficas.

Basta a nuestro propósito introductorio la presentación intuitiva de los elementos técnicos —principalmente matemáticos— de los métodos y teorías científicas que irán saliéndonos al paso. Con todo, aquellos elementos técnicos son precisamente el meollo de estos métodos y teorías, de modo que nadie que pretenda entrar en detalles podrá hacerlo sin pasar antes por los textos clásicos y las obras de consulta referenciadas en el cuerpo del texto y recogidas en la sección bibliográfica.

La riqueza del terreno que recorren estas páginas excede con mucho las posibilidades de un librito de esta extensión. En nuestra

trayectoria por la historia y la actualidad de las ciencias cognitivas habremos de contentarnos con una aproximación divulgativa, y en el peor de los casos con la mera alusión. Nuevamente, las referencias allanan muchos caminos en los que no podremos detenernos.

Hablábamos de abordar tareas «propriadamente filosóficas», y aunque la expresión es difícil de desenredar —particularmente aquí, pues hay pocas empresas científicas tan empapadas de filosofía en cada una de sus capas como las ciencias cognitivas—, precisemos que nos limitaremos a la más básica de esas tareas: la de pensar las ciencias cognitivas en tanto disciplina científica. Esa tarea puede encontrarse ya implícita en buena parte de los textos clásicos en ciencias cognitivas, de forma que volver sobre ellos equivale a acercarse a los primeros balbuceos de la filosofía de las ciencias cognitivas —balbuceos, por lo general, no superados aún en claridad y feracidad.

El itinerario que seguiremos puede reconstruirse sin dificultad a partir del esquema que ofrece el índice. Primero recorreremos la génesis histórica de las ciencias cognitivas (capítulo 1); nos detendremos a continuación a considerar los paradigmas computacionales clásicos y sus derivas contemporáneas (capítulo 2); comentaremos luego las principales objeciones filosóficas al computacionalismo (capítulo 3); después echaremos un vistazo a las alternativas teóricas al computacionalismo (capítulo 4) y, finalmente, nos acercaremos a la más amplia de las discusiones metodológicas en el área (capítulo 5). Será sobre todo en ese último tramo que nos alejemos de la exposición para sumergirnos brevemente en la argumentación —el cambio de registro podrá identificarse con facilidad.

Durante el curso 2009/2010 fui alumno de la asignatura de máster cuya denominación coincide con el título de este libro. Para entonces, llevaba varios años ya tomando notas de lectura de los clásicos de la psicología cognitiva, la inteligencia artificial, las neurociencias y la filosofía relacionada con ellas; hasta hoy, he mantenido esa costumbre. En el curso 2022/2023 empecé a hacerme cargo de esa asignatura en el Máster en Epistemología de las Ciencias Naturales y Sociales de la Universidad Complutense de Madrid, y emprendí entonces la tarea de organizar aquellas notas para impartirla. Este libro no es otra cosa que eso: aquellas notas organizadas con la vista puesta en la docencia. Se trata del libro que hubiera querido leer en aquel curso 2009/2010. Confío en que será útil a sucesivas promociones, y agradezco sinceramente a las previas su inestimable aportación.

A pesar de que su génesis respondiera a las señaladas intenciones pedagógicas, este libro no está exclusivamente dirigido a estudiantes de posgrado. Cualquiera puede adentrarse en sus páginas, y aunque la lectura de algunos pasajes puntuales pueda resultar exigente, en ningún caso se trata de un desafío disuasorio.

Entre la tentación de la lista interminable de agradecimientos y el apuro de la omisión, opto por la vía media de reservar estas palabras de agradecimiento para una sola persona: Luis Fernández Moreno, coordinador del máster en el momento de mi incorporación. Aprovecho pues este huequecito para expresarle mi agradecimiento por su generosidad y su ejemplo, no sólo de integridad personal y rigor académico, sino asimismo de compromiso institucional auténtico, desinteresado.

# 1

## La revolución cognitiva

### 1. Introducción

Antes de entrar a comentar los orígenes históricos de las ciencias cognitivas convendrá acotar tentativamente de qué hablamos cuando hablamos de ciencias cognitivas. Al efecto, cabe comenzar por destacar que las ciencias cognitivas constituyen una empresa de marcada vocación interdisciplinar en la que conviven diferentes «paradigmas», diferentes formas de entender y practicar las ciencias cognitivas.<sup>1</sup> La orientación programática original se formuló en términos explícitamente computacionales: la mente es un sistema que manipula representaciones realizando cálculos sobre ellas, y las ciencias cognitivas no serían otra cosa que el proyecto científico destinado a desentrañar la arquitectura computacional de la mente y la estructura computacional de las diferentes «facultades» que la integran —percepción, atención, memoria, lenguaje, etc.

Como apuntábamos, existen actualmente diferentes paradigmas en ciencias cognitivas, y algunos de ellos han abandonado explícitamente el modelo computacional en que se basó inicialmente la interdisciplina. Cada uno de esos paradigmas ofrece distintas aproximaciones al carácter y el objeto de las ciencias cognitivas y,

---

1 En su acepción habitual en filosofía de la ciencia, un paradigma es un conjunto de creencias, valores y técnicas compartidas, un «modelo o patrón aceptado» por una determinada «comunidad científica» acerca del modo de entender su disciplina y trabajar dentro de ella.



así, aunque se hable con frecuencia de *ciencia cognitiva* en singular, resulta más apropiado hablar de *ciencias cognitivas*, en plural, dada no sólo la convivencia de distintos paradigmas, sino sobre todo la aún escasa integración de las especialidades que convergen en la interdisciplina.<sup>2</sup>

Tradicionalmente, se ha incluido entre las disciplinas cognitivas a la psicología, la inteligencia artificial (IA), la filosofía, la lingüística, las neurociencias y la antropología.<sup>3</sup> Cada una de estas disciplinas ha jugado un papel distinto y cambiante en el desarrollo de las ciencias cognitivas. En el núcleo de la interdisciplina se ubicaron desde el principio la psicología y la IA, predominando inicialmente esta segunda para ceder después el testigo a la primera (Houdé, 2004: intro.; Gentner, 2010; Núñez et al., 2019). Por otra parte, el influjo de alguna de aquellas disciplinas en la empresa común de las ciencias cognitivas ha sido ciertamente irregular; en algunos casos, como el de la antropología, ha tendido a disminuir, siendo así que importantes investigadores «no la reconocen hoy como una de las ciencias cognitivas, o tan siquiera la consideran como una interlocutora

---

2 Algunos han creído escuchar en la elección entre el uso en plural o singular ecos de la polémica en torno a la unidad de la ciencia (González, 2008). Sea como fuere, y aunque el uso en plural recoge con mayor realismo la heterogeneidad del marco teórico, metodológico e institucional en el que han venido desarrollándose las ciencias cognitivas (Von Eckardt, 1993: 1; Martínez-Freire, 1992a; 2001a; Miller, 2003; Núñez et al., 2019), los hay que optan por la forma singular con la intención de subrayar la necesidad de avanzar hacia la cohesión e integración de una verdadera interdisciplinaridad (Boden, 2006: 12; Martínez-Freire, 2001a; 2001b: § 4; Bermúdez, 2014: 88 y ss.). De hecho, el uso original en singular apuntaba en esa dirección, y no faltan los que sostienen que cuanto quedaría hoy sería constatar el fracaso de ese intento de integración (Serrano, del Castillo & Carretero, 2014; Núñez et al., 2019; cf. Gentner, 2019). Anotemos, como anécdota histórica, que el uso en plural antecedió (Longuet-Higgins, 1973: 37) al primer uso en singular (Norman et al., 1975: 409).

3 Aunque la frontera puede resultar en ocasiones difusa (Boden, 2006: 685), es necesario diferenciar entre la IA entendida como disciplina científica dedicada al estudio de capacidades mentales (Boden, 1977: 23) y la IA entendida como rama de la ingeniería dedicada al desarrollo de sistemas computacionales que imitan ciertos aspectos de la cognición humana con fines exclusivamente instrumentales. Salvo indicación en sentido contrario, empleamos en lo sucesivo la locución en su primera acepción.

potencialmente interesante» (Beller, Bender & Medin, 2012: 3; v. et. Bender, Hutchins & Medin, 2010).

Todas las áreas de investigación que integran las ciencias cognitivas comparten, claro, uno y el mismo objeto: «el estudio científico multidisciplinario de la cognición» (Bechtel, Abrahamsen & Graham, 1999: 3), según la fórmula habitual en la literatura. No obstante, lo cierto es que no existe nada parecido a una definición consensuada de «cognición», y si bien «algunas definiciones pueden considerarse mejores que otras, es ciertamente improbable que ninguna definición pueda cubrir todos los usos legítimos del término: cualquier definición de ‘cognición’ implicará un cierto grado de estipulación» (Bayne et al., 2019: 608).<sup>4</sup> De este modo, los partidarios de cada uno de los paradigmas que conviven en ciencias cognitivas definen de forma diferente la cognición: simplificando, mientras los computacionalistas la conciben en términos de procesamiento de información vía manipulación interna de representaciones, los anticomputacionalistas nos la presentan como el control por parte de los organismos de sus acciones en sus medios (Dawson, Dupuis & Wilson, 2010: 39). De cara a esbozar, con todo, una definición tentativa y en cierta medida neutral, podemos apelar a aquello que de forma más radical distingue al reino animal del resto de la naturaleza: la capacidad para recordar el pasado, escoger en el presente y planear para el futuro (Wasserman & Zentall, 2012: 2).

Como sugeríamos, diversas áreas de las neurociencias, la psicología, la lingüística, la antropología, la filosofía y la IA participan en este proyecto de aproximación científica a la cognición. El momento en que comenzó a fraguarse la convergencia entre estas áreas ha sido fechado con mucha precisión: el 11 de septiembre de 1956. Aquél fue el día en que se celebró la segunda jornada del *MIT Symposium on Information Theory*.<sup>5</sup> En ella, George Miller

---

4 George Miller, uno de los protagonistas de la revolución cognitiva (vid. *infra*, § 2), señala respecto del uso inicial del término en este contexto: «nuestra intención al utilizar la palabra ‘cognición’ era la de marcar las distancias con el conductismo. Buscábamos una etiqueta que remitiera a lo mental, pero ‘psicología mental’ sonaba terriblemente redundante» (Miller, 1986: 210).

5 Fueron Newell y Simon (1972: 4) quienes fecharon el nacimiento de las ciencias cognitivas de forma tan precisa. Esta datación se popularizaría unos años después al tomarla Gardner prestada de una ponencia de Miller de junio de 1979 (Miller,

presentó su clásico «El mágico número siete», demostrando la existencia de unos concretísimos límites de la memoria a corto plazo, en torno a siete ítems; Chomsky, sus «Tres modelos para la descripción del lenguaje», avanzando la concepción de la gramática como sistema generativo y exponiendo el núcleo de las ideas que desarrollaría en *Estructuras sintácticas*, de 1957, obra fundacional de la lingüística teórica contemporánea; y Allen Newell y Herbert Simon, «La máquina de la teoría lógica», haciendo públicas las primeras demostraciones de teoremas de los *Principia Mathematica* de Whitehead y Russell realizadas por un programa de ordenador, en agosto de aquel mismo año. Comentando posteriormente sus impresiones, Miller apuntaría que abandonó ese día el simposio con la convicción de que las áreas que en él confluyeron «formaban parte de un todo mayor y en el futuro experimentarían una elaboración y coordinación progresivas de sus preocupaciones compartidas» (Miller, 1979; 2003: 143).

Aunque esta pretensión de precisión en la datación del nacimiento de las ciencias cognitivas pueda sonar a chiste, lo cierto es que en ese año de 1956 tuvo lugar una importante cantidad de acontecimientos de gran relevancia para la configuración del marco en el que a partir de ese momento comenzarían a estudiarse los fenómenos mentales. Así, por ejemplo, Jerome Bruner, Jacqueline Goodnow y George Austin publicaban ese año *A Study of Thinking*, texto que podemos leer hoy como un punto de transición entre la psicología del *New Look* y la del procesamiento de información; Claude Shannon y John McCarthy editaban la célebre compilación *Automata Studies*; y Newell y Simon, junto con Marvin Minsky, Nathaniel Rochester, Warren McCulloch, Oliver Selfridge, McCarthy y Shannon, se reunían en la *Conferencia de Dartmouth*, considerada actualmente como el evento fundacional de la IA.

Las ciencias cognitivas son «hijas de los cincuenta» (Miller, 2003: 141): nadie se refería aún a ellas por su nombre, pero existían ya entonces como un proyecto con entidad e identidad propias, y con expresa vocación interdisciplinar. En cualquier caso, para comprender su génesis tendremos que retroceder algunos pasos y remontarnos unas cuantas décadas atrás.

---

1979) e introducirla en su historia de la revolución cognitiva (Gardner, 1985: 28; v. et. Miller, 2003).

## 2. La revolución cognitiva

Siendo los problemas que hoy abordan las ciencias cognitivas los que han ocupado a la psicología desde sus orígenes, ésta constituye el centro de gravedad de la interdisciplina. Puede ilustrarse esta idea con el siguiente ejemplo. Desde la fundación de la revista *Cognitive Science*, en 1977, las áreas de especialidad de la mayoría de los autores que publican en la misma han sido la psicología y la IA. Durante los primeros años de vida de la revista predominaron los especialistas en IA, pero ese predominio pronto se convertiría en equilibrio y, así, al término de la primera década de la trayectoria de la revista, psicólogos y especialistas en IA se repartían equitativamente dos tercios del espacio en la misma. La proporción de psicólogos continuó aumentando, y actualmente más de la mitad de los autores que publican en *Cognitive Science* son psicólogos (Gentner, 2010; Núñez et al., 2019). Ha llegado incluso a sostenerse que existe un solapamiento casi total entre la psicología actual y las ciencias cognitivas (Anderson, 2009: 10; Mandler, 2007: 219).

Dada esta centralidad de la psicología en las ciencias cognitivas, un buen modo de aproximarse al nacimiento de la interdisciplina consiste en atender al cambio de paradigma que se produjera a mediados del siglo XX en psicología, del conductismo al cognitivismo. Para comprender este cambio de paradigma debemos dar el referido par de pasos atrás.

En Wilhelm Wundt suele reconocerse al padre de la psicología científica, tanto por el espíritu sistemático de su obra como por su ímproba labor en la instauración de un contexto institucional para el desarrollo de la disciplina —destaca en este sentido la fundación bajo su dirección del primer laboratorio de psicología experimental en la Universidad de Leipzig, en 1879.

La escuela fundada por Wundt tuvo un claro carácter introspeccionista. Consideraba que los estados mentales conscientes —la «experiencia inmediata», en sus términos— constituyen el objeto de estudio de la psicología. Para desentrañar ese objeto se hacía necesario recurrir a la introspección, dado que nadie más que el propio sujeto de la experiencia tiene un acceso directo a ella. El psicólogo experimental, por su parte, sólo dispone de un acceso indirecto a los estados mentales de los sujetos experimentales, mediado por los informes que éstos emiten al ser expuestos a diferentes situaciones experimentales.