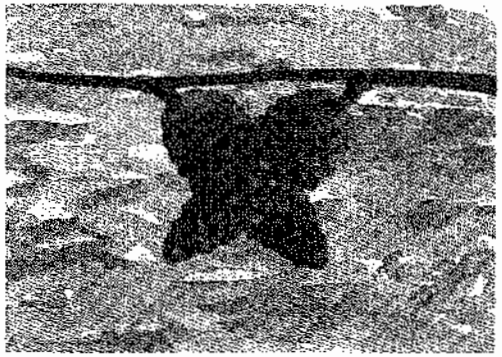
Abril 1991 N.º 119 / 700 ptas.

Revista de Occidente

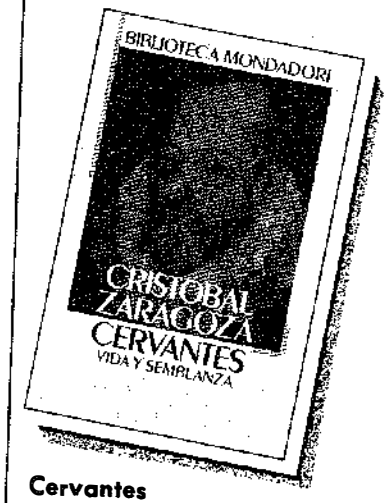


LA MENTE Y SUS REPRESENTACIONES: COGNICION E INTELIGENCIA ARTIFICIAL

VIOLETA DEMONTE: *De la complejidad del «logos». Y a modo de presentación* ● JOSÉ E. GARCÍA-ALBEA: *Entrevista con Jerry Fodor* ● JOSÉ E. GARCÍA-ALBEA: *La mente como máquina simbólica* ● JUAN A. GARCÍA MADRUGA: *El enfoque computacional en el estudio del desarrollo de la mente* ● PHILIP N. JOHNSON-LAIRD Y RUTH M. J. BYRNE: *Modelos mentales en el razonamiento deductivo* ● TERRY WINOGRAD: *Máquinas pensantes: ¿son posibles? ¿lo somos?* ● RAFAEL PARDO AVELLANEDA: *Sociología y ciencia cognitiva* ● ROGER PENROSE: *¿Para qué sirve la mente?* ● PABLO DEL BARCO: *Un poema.*

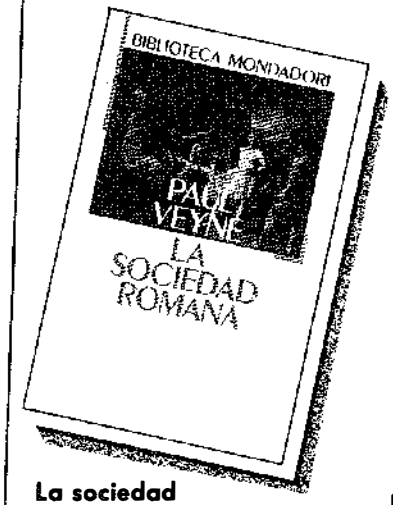
Viñeta: José Manuel Calzada

BIBLIOTECA MONDADORI

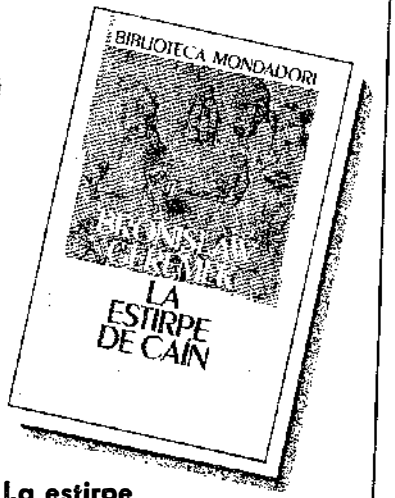


Cervantes
Vida y semblanza
CRISTÓBAL ZARAGOZA

Figuras helénicas
y géneros literarios
CARLOS GARCÍA GUAL



La sociedad
Romana
PAUL VEYNE



La estirpe
de Caín
BRONISŁAW GEREMEK

De la complejidad del 'logos'. Y a modo de presentación

Violeta Demonte

Trescientos cincuenta años después, algunas de las distinciones casi inaugurales de René Descartes tienen la transparencia de lo que, siendo normal, resulta ser a la vez clarividente. Esa tajante diferenciación suya entre la *mente* y el *cuerpo,* lo que se vino a llamar su *dualismo,* es la primera piedra de la práctica totalidad de una futura ciencia del conocimiento humano; sin ir más lejos, la historia toda de la psicología, como seguramente se habrá dicho mil veces, es poco más que un conjunto de reacciones a su idea directriz: «... de manera que ese yo, es decir el alma (*âme* en francés, *mens* en el texto latino), por la cual yo soy lo que soy, es enteramente distinto del cuerpo, e incluso más fácil de conocer que éste» (*Discurso del método.* Cuarta Parte).

Pero repárese en la disquisición de fondo que precede a esa propuesta fundacional: «Examinando atentamente lo que yo era, y viendo que podía fingir que no tenía cuerpo alguno..., pero que no por eso podía fingir que yo no fuese; sino que, por el contrario, del hecho mismo de que yo atinase a dudar de la verdad de las demás cosas se

seguía de una manera muy cierta y evidente que yo era; mientras que si hubiese dejado de pensar, aunque todo el resto de lo que imaginaba hubiese sido verdad, no tendría razón alguna para creer que yo fuese; conocí con eso que yo era una sustancia cuya esencia o naturaleza toda no es sino pensar y que no necesita, para ser, de ningún lugar ni depende de ninguna cosa material» (*Discurso,* Cuarta Parte). Parecería que nos viene a decir aquí Descartes, con ojo predecesor del actual enfoque computacional, que en la explicación y caracterización científica del ser humano podría resultar por entero prescindible el hecho de que seamos sistemas biológicos.

La deuda con el cartesianismo es también lo que concierne al atisbo de la metáfora del hombre como máquina, a la posibilidad de simularlo por medio de artefactos y a la imposibilidad, sin embargo, de así representar lo que viene a ser específicamente humano; sobre todo, la de simularlo cabalmente: «... si hubiese semejantes máquinas que tuviesen los órganos y la figura exterior de un simio o de algún animal desprovisto de razón no tendríamos medio alguno para reconocer que no son en todo de la misma naturaleza que estos animales; por el contrario, si hubiese otras con la forma exterior de nuestros cuerpos y que imitasen nuestras acciones... dispondríamos de dos medios muy seguros para reconocer que no por eso serían verdaderos hombres. El primero es que nunca podrían usar palabras y otros signos componiéndolas como lo hacemos nosotros para expresarles a los otros nuestros pensamientos... Y el segundo es que aun cuando hiciesen diversas cosas tan bien o mejor que ninguno de nosotros, fallarían inevitablemente en algunas otras, por las cuales se descubriría que no obraban por conocimiento sino solamente por la disposición de sus órganos» (*Discurso,* Quinta Parte). Si nos ponemos en plan exegético y desmesuradamente interpretativo podríamos exagerar diciendo que Descartes nos anticipa ya, en estas líneas, el fracaso ante el *test* sugerido por Turing en 1950 (desde una perspectiva operacional cabe decir

que una máquina piensa si no podemos distinguir su actuación de la de un ser humano ante una tarea determinada), la pregunta chomskiana sobre la razón de ser de la infinita creatividad en el uso del lenguaje, y los dilemas actuales de los científicos de la inteligencia artificial [«los hombres son fines en sí mismos y no máquinas», dirían los antimaterialistas de hoy; «la simulación de lo que pueda considerarse como genuinamente inteligente es algo aún muy lejano», afirmará Penrose *(La nueva mente del emperador)*]. Si el plan es crítico, nos convendrá recordar que, al adherirse nuestro filósofo a la noción griega de 'logos' como unidad inseparable de lenguaje ('palabra', 'argumento') y cognición ('razón'), estaba con ello enmascarando una relación mucho más enrevesada de la que por aquellos años podía concebir un ser reflexivo.

No es Descartes o el cartesianismo, sin embargo, el asunto de estas líneas y de este volumen sino, acaso, la rehabilitación, reconstrucción, auge y extensión de ideas como éstas en el momento actual. La cuestión es del todo pertinente: desde hace ya aproximadamente cuarenta años científicos de varias disciplinas han convergido, con ángulos investigadores, densidad teórica y medios instrumentales diversos, en el estudio de la mente. El resultado académico es una nueva disciplina bastante real y trabada (si bien institucionalmente inexistente en medios universitarios atrasados como viene a ser el español) que se suele denominar *ciencia cognitiva.* El resultado intelectual es que sabemos ahora con algo más de precisión —sin tener aún demasiadas certezas— qué es el lenguaje, cómo funciona el conocimiento, cómo podría ser posible el pensamiento o cuál es la relación entre las representaciones mentales y las inferencias que los seres humanos somos capaces de realizar. Y hasta parece que podríamos empezar a intuir cómo ciertos contenidos de los estados mentales pasan a convertirse en creencias culturales.

El asunto, como decía, es de inusual interdisciplina-

riedad en un momento en que, hasta en el saber, todo es restricción y acotación de metas y funciones; y aquí su primera y no menor virtud. El estudio de la mente, en efecto —su estructura, naturaleza y contenido; sus procesos y modo de operar; el intento de reproducirla y de simular la conducta inteligente— es hoy de interés central para varias zonas de las ciencias del hombre y de la conducta. Los filósofos de la mente reflexionan sobre la relación entre el lenguaje, el pensamiento y la realidad; si el meollo de esa filosofía es, para decirlo con palabras de Fodor, «la cuestión metafísica del lugar del significado en el orden del mundo», habrá que deslindar entonces el contenido simbólico y representacional de un hipotético «lenguaje del pensamiento». Para los psicólogos cognitivos esa pregunta se tornará en inquisición sobre los mecanismos; y se elaborarán, pongamos, modelos diversos (conexionistas, computacionales, neoestructuralistas) del razonamiento o del procesamiento de las diferentes clases de información. La lingüística, por su parte, viene dejando claro que existe algo así como una facultad del lenguaje que parece ser independiente de los otros sistemas cognitivos; para Chomsky se trata de un dispositivo innato y preespecificado que sólo requerirá el apropiado impulso del entorno para ponerse en marcha. Aunque generalmente aceptada en su esencia programática, esta concepción revolverá a los psicólogos que saben un poco sobre la «construcción» del conocimiento y sobre los límites —precisos pero no obvios— entre lo social y lo psicológico. En todo caso, y en cuanto a esto último, etólogos, antropólogos o pragmatistas (Sperber y Wilson son un buen ejemplo y configuran todo un programa de trabajo, aunque no una elaboración teórica cabal) se incorporan a la empresa cognitivista coincidiendo en que la conducta social y comunicativa tiene profundas raíces innatas; y en querer hacer ver también que las prácticas culturales podrían poseer una estructura conceptual relativamente uniforme. Para los neurofisiólogos la tarea será más de corte experimental y se tratará, por

ejemplo, de precisar la base material —celular— de los complejos procesos perceptivos y motores.

Y luego está la nada nimia cuestión de la aplicación, de la virtualidad social y material de lo que se teoriza y se descubre; por aquí es por donde viene la expresión deslumbradora y deslumbrante de «inteligencia artificial»: el traje nuevo del emperador sobre el que ironizara el célebre matemático antes mencionado. Los científicos de la computación llevan muchos años (desde comienzos de los cincuenta) intentando imitar la actividad mental humana por medio de sistemas «inteligentes». Estos sistemas simulan y ejecutan actividades de utilidad práctica, por lo general excesivamente rápidas y complicadas para ser llevadas a cabo por seres humanos en solitario —quehaceres «expertos»— pero también tareas cognitivas, de comunicación, de procesamiento o de solución de problemas. Los resultados son abundantes (con independencia de su sustantividad) y hay muchos grupos abocados a tal empresa; la valoración en profundidad de aquéllos, sin embargo, parece confirmar por el momento los augurios cautelosos de las almas escépticas: la inteligencia artificial podría no haber llegado más allá de la réplica bien de lo que el cerebro humano puede hacer sin «esfuerzo» alguno, bien de tareas cuantitativas pero no cualitativamente complejas (la construcción de elaborados sistemas para problemas relativamente triviales). En todo caso, cualquier versión «fuerte» (capaz de concebir un cerebro meramente algorítmico) de la IA resulta, hasta ahora, de escasa plausibilidad. Uno de los pioneros en el campo, Terry Winograd, nos aproximará luego a este escepticismo con algo más de rigor. Pero conviene reparar asimismo en que aquí como en otros terrenos se reproduce la dialéctica del amo y el esclavo, y la aplicación puede terminar —en alguna medida— predeterminando la teorización. Mejor dicho, la aproximación computacional a los fenómenos del pensamiento también ha permitido formular con mayor precisión los problemas que se aspira a resolver. Ha llevado tal vez a entender

mejor cuál es el lugar de la mente en esa procelosa relación entre lo corpóreo y lo inmaterial: he aquí cómo lo dice Johnson-Laird: «Antes de la computación existía una tajante distinción entre el cerebro y la mente: el primero era un órgano físico, el segundo, una fantasmal no-entidad, apenas concebible como respetable asunto de investigación... Después de los computadores, ese escepticismo no ha ya lugar: una máquina puede ser controlada por un 'programa' de instrucciones simbólicas, y nada hay de fantasmal en un programa de ordenador. Acaso la mente es al cerebro lo que un programa es a un computador» (*The Computer and the Mind.* Prólogo).

Pero la conveniencia de la interdisciplinariedad, del intento de superar la tediosa división de la realidad en un conjunto de problemas sólo arbitrariamente independientes, no acaba aquí. Una consecuencia grata de ella es que hace posible que los científicos se cambien de chaqueta; en este caso para bien y no para medrar. Los lingüistas se pasan a filósofos; los filósofos del lenguaje se convierten en filósofos de la mente; los ingenieros devienen cuasi-psicólogos y se abocan a precisar los mínimos elementos que intervienen en la percepción del lenguaje; los físicos y matemáticos reflexionan sobre la conciencia; los biólogos investigan sobre la representación de las relaciones espaciales, por ejemplo. Tras décadas de cortes epistemológicos y criterios de demarcación, la idea de regreso a una hipotética ciencia total, aristotélica, genera una cierta euforia descansada. Invita también a la sensación de elitismo: esto no pasa ciertamente en los cerrados campos histórico-filológicos donde toda salida del corsé ficheril y pedestremente 'metodológico' se toma por gala literaria y no se entiende bien. Menos sucede aún en las instrumentales disciplinas socio-normativas, cuya fragilidad conceptual y pereza ante la teoría se enmascara a veces bajo el látigo del recurso abundante a la cuantificación y a la validez inmediata —en la coyuntura— de las generalizaciones, analogías y extrapolaciones que nos suelen brindar. No conviene dejar de señalar, en todo

caso, por mor de la claridad y de la sinceridad, que la interpretación entre los resultados de las varias disciplinas que he ido mencionando es aún tenue, y más materia de libros de 'lecturas' o de congresos ambiciosamente concebidos que de proyectos específicos o de teorías verdaderamente globales; aunque los pasos iniciales estén dados.

Probablemente todo lo que he dicho hasta aquí sea lo de menos, sea el *glamour* de nuestra historia (acaso sólo interesante para quienes ven el mundo en términos de buenos y malos, tontos y listos, fascinantes o repudiables). Lo de más, como siempre, es lo que ella de verdad contiene, lo que ha permitido comprender mejor, lo que se entiende a medias y provoca disputas, lo que se infiere y se puede pronosticar a partir de lo que se dice. Este número de la *Revista de Occidente* nace no sólo del deseo de contribuir a hacer conocer un campo atrayente y abrumador; anhela que pueda permitir asomarse —aunque sea de puntillas y con dificultad— al interior del mismo, al tipo de disquisición teórica y a la clase de problemas que le conciernen a la ciencia de la cognición. No ofrece una perspectiva completa de la materia porque ello habría requerido, entre otras muchísimas cosas, hacer entrar también a la neurociencia (quizá de lenguaje más inaccesible para nuestros lectores y lectoras y por ello descartada de antemano), pero sí una perspectiva —creo— bastante pertinente (o relevante), para usar el término tan caro a varios de los especialistas que aquí escriben. Se me permitirá ahora, para así cerrar esta presentación, evocar críticamente algunas de las consideraciones de las páginas que vienen e intentar introducir a los lectores en una especie de guía interesada a la lectura de este número.

Jerry Fodor, en la entrevista con García-Albea, abre a medias varias de las ventanas que nos acercan a los problemas medulares: la naturaleza de la ciencia cognitiva y su lugar entre las ciencias básicas y las especiales, los diversos enfoques que podrían orientar su programa

de trabajo y la condición intrínseca ('modular' o no modular) de los procesos de los que debe dar razón. La ubicua noción de 'funcionalismo' es a su juicio la adecuada para sentar un nivel explicativo independiente para la ciencia de la mente: «un nivel de representación teórica que se aplica a los organismos biológicos pero no está especificado en términos biológicos». De ahí a la idea de que las actividades mentales son computacionales hay un paso muy corto (si media la convicción de que existen representaciones mentales casualmente conectadas), por lo menos desde el ángulo epistemológico en que ese autor los sitúa. De ahí también la crítica final de Fodor al 'conexionismo' (que el lector puede comparar con la de Winograd a esa 'inteligencia emergente') por la concepción que entraña de los procesos mentales: éstos serían finitos y esencialmente no productivos. Pero las propuestas de Fodor no son sólo programáticas, la distinción clave entre procesos cognitivos aislados —no globales, 'impenetrables' o modulares— ha sido empíricamente muy fructífera para el estudio del lenguaje, de la visión o de ciertas áreas de la percepción. Ahora bien, si el razonamiento es no-modular, como concluye este filósofo, resultará que no sabemos bien cómo acceder a un contenido central del paradigma cognitivista; pero ya volveré sobre esta cuestión. El nítido ensayo de García-Albea traza la perspectiva histórica de la idea del hombre como máquina y razona que la concepción actual de éste como máquina simbólica —que es la fodoriana y, a la postre, la de Newell y Simon— surge de un compromiso entre el dualismo y el mecanicismo. El argumento de este autor, al igual que el del precedente, es sobre todo apriorístico: hay estrecha conexión entre símbolo e intencionalidad, que es la propiedad característica de la mente; se añadirán aquí, no obstante, otras consecuencias que convertirían a esta concepción también en imprescindible desde el punto de vista de los hechos. La presentación del enfoque computacional (con lo que parece ser la constante velada objeción al cuasi antagónico conexionismo) se completa en el inteligente estudio crítico de García Madruga. Una idea medular de su disquisición es que los modelos computacionales son capaces también de dar cuenta del 'desarrollo' intelectual (cuestión que quedaría fuera de sus objetivos iniciales), en tanto en cuanto estos sistemas incorporen estrategias de aprendizaje, de automodificación y de interacción con el entorno. Asunto central al ensayo de Johnson-Laird —uno de los más notables investigadores de la mente en el momento actual— es cómo el ser humano ejercita la capacidad de hacer deducciones válidas. La tesis que viene defendiendo, crucial para la teoría de la racionalidad y que ilustra en estas páginas con tres estudios de casos, es que el razonamiento humano no es un proceso sintáctico formal (como suponen en alguna medida los sistemas computacionales), sino uno de comprensión de significados y de manejo de 'modelos' mentales del mundo. La observación de Fodor, pues, acerca del carácter del razonamiento podría ser retomada a lo largo de líneas como éstas.

Hasta aquí lo relativo a la cognición; vayamos ahora a sus efectos y a la posible contrastación, en la aplicación, de ideas como las que hemos esbozado. Quede por delante, de todos modos, que ni las dificultades de la inteligencia artificial serían un índice obvio del error del paradigma cognitivista, ni los aciertos de éste tienen por qué garantizar logros en el otro terreno. La simulación de inteligencia posee bastantes de las características de los saberes prácticos: recuérdese que la resolución de problemas ingenieriles a lo largo de la historia no ha parecido depender crucialmente del desarrollo de la física, por ejemplo; y en ello reside su cara y su cruz. Es evidente también, como en todos los casos, que el grado de refinamiento que una tarea tan ambiciosa pueda llegar a alcanzar no será independiente de los principios básicos sobre los que se asiente. El brillante trabajo de Winograd —cargado de conocimiento directo de la materia y de inusual capacidad de disección crítica— ilumina

minuciosamente los principios en los que se sustenta la construcción de máquinas pensantes, la historia de su desarrollo, las características de los frutos de la investigación sobre ellas; y hasta propone un giro quirúrgico-rectificador, acaso poco perspicuo a primera vista, como luego diré. La filosofía del proyecto general, que Winograd denomina «racionalismo (hecho) de retazos» conduce, a su modo de ver, a la caracterización de una inteligencia burocrática, rígida, obtusa e incapaz de adaptarse a situaciones nuevas. Los «sistemas expertos» que la encarnan, sistemas reglados manipuladores de símbolos, tienen, por ello, a diferencia de los sistemas reales, serias limitaciones. No pueden, por ejemplo, relacionar representaciones actuales con representaciones previas, son ciegos para interpretar los nuevos términos o reglas que surjan en las situaciones concretas y sólo pueden funcionar en dominios especializados y muy constreñidos. A una empresa así concebida Winograd le enfrenta la alternativa de un «constructivismo hermenéutico», fundado en la concepción del ordenador como instrumento de comunicación y en la idea gadameriana de que el lenguaje es en esencia «metafórico, vago e indefinible». La propuesta es breve y muy general —aunque fascinante— y suscita el encanto y el azoro de lo que se presenta como un amor a primera vista. Habrá que ver cómo se las arregla esta nueva programación con las dificultades inherentes a la formalización explícita de los actos de habla, o cómo integra en su formulación otras concepciones de lo metafórico en tanto que fenómeno también cognitivamente determinado (la de Sperber y Wilson, pongamos); por no decir que parece estar impregnada de una distinción no clara entre lo que es un proceso lingüístico y un proceso de pensamiento en sentido estricto. En todo caso, lo que este colofón afirma de manera inequívoca es que resulta imprescindible buscar nuevas vías de enfoque; y parece haberlas. La discusión incisiva y reflexiva de Pardo tiene su origen en el cambio de estilo de los científicos sociales que se preocu-

pan por la construcción social de los sistemas tecnológicos —asumida, eso sí, críticamente— y, en especial, por las variables psico-sociales que deben intervenir en el diseño de tecnologías complejas. Sus consideraciones sobre dos áreas en que esa intervención se verifica de manera candente: los estudios acerca de la interacción ser humano-computador y la posibilidad de considerar que no es el cerebro sino la organización social una metáfora plausible de los sistemas inteligentes, esbozan un proyecto sustantivo de giro psico-social y organizativo de la ciencia cognitiva y de la inteligencia artificial. Penrose, en clave literario-filosófica —quizá como entreacto de un libro por lo demás muy técnico—, sitúa la cuestión de la inteligencia en el debate físico y metafísico de la naturaleza y el lugar de la conciencia. Algo más lejos de las cuestiones generales, en el borde de ellas, pero con fuerte interés preteórico y especulativo, se quedan las consideraciones de Sperber sobre los tipos de creencias y su relación tanto con las representaciones como con la organización cognitiva recogidas en su artículo «La epidemiología de las creencias», que no ha podido incorporarse a este número por razones de espacio y que *Revista de Occidente* publicará próximamente.

Es posible que lo trabajos que aquí recogemos no delimiten del todo el territorio de la ciencia cognitiva; es bastante seguro que no esclarecen todos los condicionamientos que a ella le impone la naturaleza misma del objeto de investigación; es más claro aún que dejan pendientes muchas cuestiones: una es el de su verdadero rigor explicativo, otra no menos decisiva la de si las varias disciplinas que la conforman llegarán alguna vez a integrarse, y cómo y dónde. El problema está ahí y viene, como diría Borges, de cuando se es «demasiado distinto y demasiado parecido»; no podrán acaso engañarse, «lo cual hace difícil el diálogo».

V.D.