

Perspectivas actuales en la filosofía de lo mental¹

1. Introducción

La discusión acerca de la mente y el cerebro hunde sus raíces en la filosofía griega, en la escolástica medieval y en el cartesianismo racionalista de los siglos XVII y XVIII, pero se expresa con inusitado interés en los más recientes estudios científicos neurobiológicos y en la subsiguiente reflexión filosófica acerca de ellos. En la actualidad, uno de los temas que más atrae la atención de los investigadores en neurociencia, psicología y filosofía de lo mental es precisamente la naturaleza de la consciencia. Un símbolo emblemático de esto es el reciente *The Nature of Consciousness. Philosophical Debates*² o lo que escribía el conocido neurólogo Antonio R. Damasio en un número especial de la revista *Time* sobre investigación: "La consciencia no es un misterio cualquiera; es el misterio de nuestro tiempo, un enigma complejo y aparentemente impenetrable. (...) Irónicamente, la maquinaria que está detrás de nuestra consciencia es *in-consciente*. (...) Sólo llegamos a conocer los resultados de esos procesos"³.

A comienzos del siglo XXI todavía no sabemos cómo pensamos, ni qué es pensar, ni tenemos ninguna teoría que lo explique satisfactoriamente. No hay una respuesta a la pregunta de si la mente es una propiedad emergente de la actividad eléctrica y metabólica del cerebro, y en el caso de que eso fuera así, no se conoce tampoco de qué forma la red nerviosa cerebral da origen a los fenómenos mentales. "Estamos apenas comenzando a ver –escribía Dennett a propósito del libro de Damasio *Descartes' Error*– los primeros puentes que se

1. Agradezco vivamente la invitación del profesor Eudaldo Forment a colaborar en *Espíritu* con este texto que ve aquí la luz por primera vez. Presenté versiones precedentes de este trabajo en la I Jornada de Filosofía de la Ciencia sobre Mente y Cerebro, organizada por el C. M. Castilla de Madrid en noviembre de 1993, en un seminario en el Departamento de Psiquiatría de la Clínica Universitaria en 1994 y en un curso doctoral de Fundamentos de Ciencia Cognitiva en marzo de 1998. Debo gratitud en particular a Sara F. Barrena y Belén Pascual por sus correcciones y sugerencias.

2. NED BLOCK Y OWEN FLANAGAN (eds.), *The Nature of Consciousness. Philosophical Debates*, Cambridge, MA, MIT Press, 1997, 843 págs.

3. ANTONIO R. DAMASIO, "A Clear Consciousness", *Time*, Winter 1997-98, pp. 88-90.

están tendiendo sobre el gran abismo de ignorancia entre la psicología y la neuroanatomía⁴. Por estas razones, la reflexión acerca del cerebro y la mente humanas resulta un campo privilegiado para el desarrollo de un trabajo realmente multidisciplinar, en el que, fruto del esfuerzo conjunto de neurólogos, psicólogos, cognitivistas, informáticos, filósofos y teóricos de la comunicación, sea posible progresar decididamente en la comprensión de los elementos esenciales del conocimiento y la comunicación humanos.

La neurología ha hecho grandes avances en el conocimiento de la estructura y el funcionamiento del cerebro. Los estudios experimentales de los neurólogos y psicólogos en las últimas décadas han aportado muchísimos datos sobre las funciones mentales de los seres humanos y su comparación con las de los primates y otros animales superiores. Los cognitivistas –nombre que se aplica a los profesionales de la nueva Ciencia Cognitiva (o para algunos, en plural, Ciencias Cognitivas)– confían, con cierta dosis de cientismo taumatúrgico, en que una integración de las aportaciones de neurólogos y psicólogos, junto con las herramientas y los conceptos teóricos desarrollados por los informáticos y los teóricos de la comunicación, dará luz definitiva sobre qué son la información y la comunicación humanas, sobre cómo se almacena la información en las estructuras cerebrales, cómo se recupera y cómo se transmite lingüísticamente.

En la última década los mayores avances en la comprensión de los fenómenos lingüísticos han tenido su origen no tanto en las teorías o intuiciones de personas singulares excepcionalmente capacitadas, sino sobre todo en el riguroso trabajo experimental, en el concienzudo estudio de las diversas patologías lingüísticas causadas por lesiones cerebrales (trastornos específicos del lenguaje: afasia, disfasia de desarrollo, etc.) o de los procesos de adquisición de lenguaje en circunstancias extraordinarias (autismo, psicosis, esquizofrenia, demencias). Los datos disponibles son cada vez más coincidentes en todas las investigaciones y permiten entrever una efectiva unificación de las investigaciones de lingüistas, neurólogos, psicólogos y filósofos. Por ejemplo, es bien conocido el caso de Christopher: se trata de un varón de alrededor de 30 años que tiene un coeficiente de inteligencia muy limitado, hasta el punto de que no es capaz de atarse los zapatos y ha de vivir permanentemente en una institución, pero que al mismo tiempo habla otras quince lenguas además del inglés⁵.

Una de las cosas que a quienes nos dedicamos a la filosofía más nos llama la atención es que algunos de los que han dedicado mucho tiempo a la investigación empírica reclaman la re-introducción de la noción de *naturaleza humana* para poder dar cuenta legítimamente de la universalidad de los resultados experimentales. Así lo hace Steven Pinker en el último capítulo de *El instinto lingüístico* (1995), o Mehler y Dupoux en *Nacer sabiendo* (1992): "Lo que este libro intenta demostrar –afirman estos autores– es que la idea de *naturaleza humana* tendría que ser el hilo conductor de la investigación en las ciencias cognitivas ". Y añaden poco más adelante: "sostener la existencia de una naturaleza

4. DANIEL DENNETT, "Our Vegetative Soul. The Search of a Reliable Model of the Human Self", *Times Literary Supplement*, 25 agosto 1995, p. 3.

5. MARTIN DAVIES, "Christopher's Languages". *Times Literary Supplement*, 25 agosto 1995, p. 8.

humana no es empobrecer al hombre ni reducir los individuos a una estepa seca y aburrida. Constituye más bien una oportunidad de determinar por fin lo que somos".

2. *El marco de la discusión contemporánea sobre mente y cerebro*

Después de esta relativamente amplia introducción, quiero ya avanzar la tesis primera de estas páginas, que puede formularse así: el llamado problema mente-cerebro no es en modo alguno el problema alma-cuerpo⁶. Entender esto es importante: no lo entendió el Nobel Francis Crick que publicaba en 1994 el libro *La búsqueda científica del alma* en el que –como escribía el profesor Mariano Artigas– "mezcla interesantes perspectivas científicas con un materialismo barato y anti-religioso, impropio de un premio Nobel"⁷. En la cultura occidental la comprensión de la conjunción de cuerpo y alma en el ser humano ha adoptado múltiples formas en los tres último siglos: dicho *grosso modo* quienes quisieron comprender esa articulación se dividieron entre dualistas –alma y cuerpo son dos elementos radicalmente distintos y de difícil combinación coyuntural– y monistas –existe sólo uno de los dos elementos– que pueden ser divididos a su vez en materialistas (solo existen cuerpos) e idealistas (que sostienen que realmente lo que existen son espíritus). Pues bien, lo primero que quiero afirmar es que el debate contemporáneo acerca de las relaciones entre la mente y el cerebro no tiene nada que ver (prácticamente) con aquella otra cuestión acerca de la articulación de alma y cuerpo. La salvedad del "prácticamente" se debe a que el efectivo carácter unitario del ser humano hace que todo lo humano tenga alguna relación. Por poner un ejemplo, tanto la demostración matemática del último teorema de Fermat, como la búsqueda de un nuevo fármaco contra el virus del SIDA o los periódicos que han salido a la calle esta mañana no tienen tampoco (prácticamente) nada que ver con la discusión filosófica acerca de alma y cuerpo. Lo que quiero afirmar con esto es que el estudio de la mente y el cerebro no es de suyo un problema metafísico, ni mucho menos religioso. No hay en el cerebro de cada ser humano un homúnculo, un hombrecillo, al que llamamos "alma", que gobierna el funcionamiento cerebral. Ese es un mito de cuño cartesiano que a fin de cuentas no explica nada, puesto que la mente no existe ni opera sin el cuerpo.

El estudio de la mente y el cerebro es un campo, o quizá una selva, en el que se han producido avances importantes en los últimos años mediante nuevas técnicas de análisis y detección de los procesos cerebrales y mediante nuevos conceptos teóricos originados en la computación y en la teoría de la comunicación que permiten explicar mejor cómo interactuamos los seres humanos entre nosotros y qué hacemos con nuestros instrumentos, las máquinas, los coches, los ordenadores. ¿Cómo funcionan nuestros cerebros, el de cada uno, y en qué medida ese funcionamiento está conformado por el medio en que cada uno ha crecido, por la educación recibida, por los estímulos externos? ¿Cuáles son los patrones o pautas básicas de la actividad inteligente humana, y en qué medida esas pautas son universales, y pueden ser enseñadas y mejoradas? Estas son –a mi juicio– las preguntas realmente importantes hoy en día que, por supuesto,

6. Debo esta formulación al profesor Jorge Vicente.

tienen alguna relación con la pregunta histórica acerca de quién es el hombre, con el problema venerable de cuál sea la naturaleza humana, pero no son de suyo cuestiones religiosas.

Quienes sostenemos una continuidad entre las diversas áreas o estratos del conocimiento humano, entre ciencia, filosofía, sentido común, religión, arte, etc. no pensamos que se traten de compartimentos estancos, pero precisamente por ello aspiramos a que las transferencias o interferencias de esos diferentes ordenes de la actividad humana sean claras, precisas y rigurosas. Lo que sí sabemos es por una parte que la respuesta acerca de lo que sea el ser humano no está oculta en las circunvoluciones cerebrales, –de modo análogo a como la comprensión de este texto no está oculta en la estructura molecular del papel o de las manchas de tinta con que están impresas mis palabras–, y por otra sabemos también que el conocimiento detallado de los diversos niveles del funcionamiento cerebral puede facilitar muchísimo nuestra comprensión de la actividad inteligente humana, considerada habitualmente como el rasgo más característico de nuestra especie.

Hechas todas estas precisiones es ya el momento de adentrarse en algunas de las corrientes actuales en la filosofía de lo mental. Concentraré especialmente mi atención en el funcionalismo, como teoría estándar más en boga tanto científica como popularmente, para pasar luego a la presentación de algunas direcciones más prometedoras en las que, en mi opinión, se encuentran las críticas más interesantes al funcionalismo. Trataré con algún detenimiento el redescubrimiento de la intencionalidad y de la subjetividad por parte de Putnam y Searle, y finalmente esbozaré las líneas maestras de la propuesta que me parece más sugestiva y prometedora, porque aspira a una comprensión más global de la inteligencia.

3. *El funcionalismo como teoría estándar y las críticas de Putnam y de Searle*

Como es conocido, a partir de 1933 con el auge del nazismo en Alemania y la II Guerra Mundial centenares de científicos, académicos y profesores europeos huyeron a los Estados Unidos. En poco más de una década estos emigrantes dinamizaron el ambiente intelectual de las Universidades americanas, dotadas ya entonces de abundantes medios económicos. Respecto de la filosofía de lo mental que aquí nos ocupa, la herencia de los científicos y filósofos del Círculo de Viena fue el *fisicalismo*, la aspiración de reducir los fenómenos psicológicos a fenómenos físicos, esto es, de explicar la mente humana científicamente mediante la *reducción* de los fenómenos psicológicos o mentales a sus características y leyes físicas. El fisicalismo encontró un buen aliado en el conductismo que se había desarrollado con anterioridad en los Estados Unidos y que en el área de la filosofía de la mente fue llamado "naturalismo". Las dos tesis básicas de esta concepción son –según Tyler Burge⁸– primero, la afirmación de que no hay nada más que las entidades identificables en las ciencias físicas o que el

7. MARIANO ARTIGAS, "El cerebro del Dr. Crick. Buscando el alma con el bisturí", en *Ace-prensa*, 156/94, 23 noviembre 1994.

8. TYLER BURGE, "Philosophy of Language and Mind: 1950-1990", *Philosophical Review* 101 (1992), pp. 32-33.

sentido común considera como físicas; y en segundo lugar, que todo el discurso mentalista habitual que habla de deseos, intenciones, creencias y aspiraciones puede ser reducido al discurso de las ciencias físicas o médicas: ésta sería la manera de *explicar* científicamente lo mental y de eliminar todos los elementos acientíficos habituales de la psicología popular.

Aunque hoy día nos resulte quizá sorprendente, el naturalismo se convirtió en una de las escasas ortodoxias de la filosofía americana a lo largo de los años sesenta. Adoptó el nombre de teoría de la identidad entre la mente y el cuerpo y aspiraba a ser una teoría empírica, diseñada conceptualmente por la filosofía, cuyos detalles serían con el tiempo completados por la ciencia. Se esperaba que su ontología incluyera sólo entidades bien acreditadas científicamente: sólo células cerebrales con su física y su bioquímica. Sus principales propulsores fueron Place, Smart y Armstrong, y sostenían que los eventos mentales (nuestras ideas, imaginaciones, recuerdos, deseos, etc.) no son ni paralelos, ni coincidentes, ni causados, ni acompañados de eventos cerebrales, sino que simple y estrictamente *son* eventos cerebrales, eventos neuronales⁹. Como es fácilmente comprensible un materialismo tan crudo –que contrasta tanto con nuestra experiencia común– suscitó controversia entre los filósofos. La objeción más demoledora provino de Hilary Putnam: si un estado mental es simplemente la estimulación de una célula o fibra cerebral –venía a decir Putnam– parece imposible que un estado mental como el dolor se identifique con un único estado neuronal en todos los organismos que sienten dolor teniendo en cuenta su diversísima fisiología. O todavía resulta menos plausible pensar que un determinado pensamiento –como el de que nueve es el triple de tres o el teorema de Pitágoras– sea el mismo estado físico –la estimulación de unas mismas neuronas– en el cerebro de todos los que lo piensan.

La respuesta a esta certera objeción fue el desarrollo de un nuevo paradigma para dar cuenta de los estados mentales en términos "científicos". En lugar de apelar a la neurofisiología cerebral, los filósofos –pero también los neurólogos– acudieron ahora a la informática como fuente de inspiración. Pensaban que identificar un estado mental con algún tipo de estado abstracto de un ordenador evitaría los problemas que suscitaba la identificación de lo mental y lo neuronal. Esta perspectiva fue denominada *funcionalismo*, porque sostiene que los estados mentales son funcionales, esto es, su singularidad es consecuencia de su papel funcional dentro del sistema cerebral completo. Me parece que ésta sigue siendo hoy en día la teoría estándar para comprender lo mental y los procesos cognitivos, avalada quizá por el impresionante auge de la informática en todos los órdenes. Quizá por esta razón el funcionalismo suele ser también denominado *teoría computacional de la mente*.

Una imagen útil para comprender el calado del funcionalismo es la de quienes sostienen que un ordenador y un ser humano a un nivel suficiente de abstracción son indistinguibles. La nueva ciencia cognitiva estudia las operaciones mentales como una manipulación de información, de símbolos, dentro del cerebro humano en términos de relaciones causales entre el cerebro, el organismo y el ambiente. La relación entre lo mental y el cerebro se piensa que ha de ser semejante a la que existe entre el flujo de información y el flujo de energía en un ordenador. Dentro del funcionalismo en boga pueden ser incluidos

tanto quienes piensan que el estudio de cómo operan los ordenadores más avanzados nos puede hacer entender mejor el funcionamiento de la mente humana, como quienes piensan que el conocimiento más detallado de cómo opera el cerebro humano nos puede ayudar a construir mejores ordenadores (Esta es la idea básica del *conexionismo* o de la construcción de "redes neuronales").

Una muestra de la vigencia del funcionalismo es –a título de ejemplo– la memoria anual del Center for the Study of Language and Information de Stanford en la que se define el área de la investigación de aquel prestigioso centro como el estudio de la actividad de los agentes inteligentes, sean estos biológicos o artificiales. A todos los que manejamos un ordenador nos resulta cuando menos sorprendente que quienes estudian con gran especialización la inteligencia consideren, por así decir, accidental o circunstancial la diferencia entre seres humanos y máquinas. La causa de esto se encuentra muy probablemente en esa concepción cientista de la objetividad a la que antes aludía –heredada del materialismo cientista de las décadas precedentes– que convierte el estudio de la inteligencia en el estudio del *comportamiento* inteligente, esto es, en el estudio de algo que es pública y científicamente observable y para lo que resulta escasamente relevante la naturaleza biológica o mecánica del agente. Una perspectiva así deja de lado cualquier intento de explicar la experiencia subjetiva que todos tenemos para otorgar primacía no a la naturaleza del sujeto sino meramente a los modos de actuar. Un ser no es inteligente porque posea inteligencia sino porque se comporta *como si* fuera inteligente.

Adoptar un punto de vista así para intentar comprender qué es la inteligencia es –en mi opinión– una manera desenfocada de mirar a las cosas. El fenómeno del enamoramiento puede servir quizá para ilustrar ese desenfoque. No hace mucho tiempo causaba gran expectación en los medios de comunicación lo que llamaban "la química del amor", esto es, la pretensión de haber descubierto finalmente que a pesar de los poetas románticos el enamoramiento consiste simplemente en una cuestión de hormonas, encimas, o determinados componentes bioquímicos corporales. Cualquiera que esté o haya estado enamorado alguna vez, se resiste con toda razón a aceptar que el enamoramiento sea meramente apetencia sexual o una extraña "sintonía hormonal" entre dos seres humanos. A mí me gusta recordar aquel dicho castizo sobre el porqué del enamoramiento, el por qué una persona concreta se enamora de otra, que explica que "a base de mirar se enciende la bombilla", aludiendo de esa manera al carácter cultural del amor, que guarda escasa relación con las hormonas. Nadie pretendería nunca comprender qué sea el amor humano mediante el estudio fisiológico, científico y riguroso de la cópula sexual, argumentando que quienes tienen una relación sexual se comportan *como si* estuvieran enamorados. Más aún me parece que casi todos pensamos que para comprender realmente qué es el amor humano, el estudio de la literatura y de las grandes creaciones amorosas es mucho más útil que la fisiología.

Acudiendo a otro ejemplo que a mí me parece todavía más ilustrativo, la pretensión de comprender la consciencia, la subjetividad o la inteligencia humana mediante el estudio físico de las neuronas cerebrales y sus conexiones podría asemejarse al intento de averiguar qué es una tarjeta de crédito mediante su concienzudo análisis al microscopio electrónico. Tal como veo yo las cosas,

sólo mediante la superación de aquel prejuicio ideológico cientista, puede lograrse un decidido avance en la comprensión de la inteligencia humana.

Los argumentos que mejor muestran los flancos más endeble del funcionalismo –su carácter reduccionista y su incapacidad de dar cuenta de la intencionalidad de la mente humana– son probablemente los elaborados en los últimos años por Hilary Putnam (Harvard) y John R. Searle (Berkeley). Estos filósofos parten de experiencias de la vida cotidiana que pueden parecer obvias, pero que no lo son para los funcionalistas¹⁰. Una de estas experiencias es la de la intencionalidad de la inteligencia humana, que no es reducible a un fenómeno observable o a un proceso de carácter meramente material. La intencionalidad –que es tanto el que hagamos algo con intención como que usemos algo físico para significar otra cosa (por ejemplo, el dibujo de un pitillo tachado para expresar la prohibición de fumar)– no es apresable, captable por quien estudia los fenómenos sólo físicamente. Aunque estudiemos el letrero de no fumar con las leyes de la física y de la química, no averiguaremos su sentido, lo que quiere decirse con él.

La intencionalidad es una propiedad que aparece indisolublemente ligada a buena parte del comportamiento humano: abarca desde la fabricación y empleo de instrumentos hasta las conductas de los seres humanos que no pueden predecir las leyes de la física ni las de la biología. La intencionalidad se manifiesta en la independencia frente a los estímulos, en la orientación a objetivos, en la productividad creativa, e incluso en la misma arbitrariedad tantas veces patente en la conducta humana¹¹. El carácter intencional de los fenómenos mentales es el que da razón de nuestra capacidad de referirnos tanto a objetos inexistentes, como a las creencias, convicciones, actitudes y opiniones de los demás, de cuya "existencia objetiva" puede no haber pruebas científicas, pero con las que están forjadas nuestra vida y experiencia ordinarias. Por ejemplo, yo puedo dejar caer un libro adrede o puede caérseme inadvertidamente. ¿Cuál es la diferencia entre ambos casos? Para cualquiera de nosotros es bastante clara la diferencia entre hacer algo sin querer o queriendo. Al menos –según me dicen– los árbitros de baloncesto creen tenerlo claro cuando deciden si una falta es intencionada y es entonces sancionada con tiro libre sin rebote y posesión de balón, o es no intencionada y en ese caso la sancionan sólo con saque de banda. Esta caracterización informal de la intencionalidad –sobre la que volveré–, puede ayudar a entender por qué se trata de un fenómeno *primitivo*, en el sentido de que no puede ser reducido a un fenómeno físico más simple ni es tampoco eliminable para comprender la inteligencia humana.

Hilary Putnam se basa en el estudio de los problemas relativos al significado de nuestras expresiones para mostrar que "el funcionalismo no funciona"¹²,

9. DANIEL DENNETT, "Perspectivas actuales en la filosofía de la mente", *Teorema* 11 (1981), pp. 204-205.

10. Sigo en líneas generales el trabajo de Pablo J. Concepción, *Reduccionismos contemporáneos acerca de la mente: La inmaterialidad de lo mental*, Tesis doctoral, Universidad de Navarra, 1994.

11. JOSÉ E. GARCÍA-ALBEA, "La mente como máquina simbólica", *Revista de Occidente* 119 (1991), pp. 47-60.

12. "So functionalism doesn't work. That is to say, it doesn't fit the phenomena". HILARY PUTNAM *Representation and Reality*, Cambridge, MA, MIT Press, 1988, p. 105.

es decir, no ofrece una explicación razonable de cómo usamos efectivamente nuestras expresiones. El funcionalismo mental (también llamado "mentalismo") identifica el significado de una palabra con algo que está en el cerebro/mente de todo hablante que sabe usar la palabra. La primera objeción es la que Putnam llama la división del trabajo lingüístico por todos bien experimentada. Mis conocimientos acerca del cerebro son muy inferiores a los de un neurólogo, pero esto no impide que cuando un neurólogo y yo hablemos de "el cerebro de N" nos estemos refiriendo a lo mismo. "El lenguaje es una forma de actividad cooperativa, no una actividad esencialmente individual. (...) La referencia se fija socialmente y no está determinada por la situación o los objetos en las mentes/cerebros individuales"¹³. Nuestras palabras, las palabras de un hablante o de una comunidad, no se refieren a un estado mental del cerebro de quien habla. Como ha insistido Putnam de modo gráfico los significados no están en nuestras cabezas. ¿Cómo van a corresponder biunívocamente a estados del cerebro en cada persona expresiones informativas tan comunes como "Hay muchos gatos en la vecindad"?

John R. Searle elige en cambio la vía del estudio de los estados subjetivos de consciencia y de los estados mentales intencionales como los deseos, las creencias y las percepciones. ¿Cómo causan exactamente la consciencia los procesos cerebrales? ¿Y cómo se explica esa causación? Esta es realmente la cuestión más importante –indica Searle¹⁴– que deben afrontar las ciencias biológicas, a pesar de que muy a menudo se rehúya esta cuestión o sea mal comprendida. Searle intenta mostrar que la consciencia y la intencionalidad son intrínsecas e ineliminables de los estados subjetivos. Contra quienes quieren explicar los fenómenos mentales sin tener en cuenta la subjetividad –como se ha procurado hacer en los últimos cincuenta años– Searle argumenta que si se elimina la subjetividad se hace imposible describir la consciencia, pues las explicaciones objetivistas presuponen una concepción de la realidad en tercera persona, objetiva, en la que no hay sitio para la subjetividad¹⁵. Por el contrario Searle piensa –a mi juicio acertadamente– que la consciencia es un fenómeno biológico, y por ello debemos pensar en la consciencia como una parte de nuestra historia biológica habitual, como la digestión o el crecimiento. Sin embargo, aunque la consciencia sea un fenómeno biológico tiene algunos rasgos importantes que otros fenómenos biológicos no tienen. El más importante es el de su "subjetividad", el de su pertenencia al sujeto como propia. Sin embargo, la subjetividad de los fenómenos mentales no excluye que se trate también de algo físico, que sean un rasgo (*feature*), una propiedad del cerebro, como la liquidez es una propiedad del agua aunque no pueda decirse de una molécula singular que sea líquida. Los niveles inferiores de los procesos neurobiológicos causan los niveles superiores como el de la consciencia, aunque no podamos señalar una neurona o una sinapsis concreta en el cerebro y decir "ésta piensa en mi abuela"¹⁶.

13. IDEM, *Representation and Reality*, p. 25.

14. JOHN R. SEARLE, "The Problem of Consciousness", *Social Research* 60 (1993), p. 3.

15. IDEM, *The Rediscovery of the Mind*, Cambridge, MA, MIT Press, 1992, pp. 99-100.

16. IDEM, "The Problem of Consciousness", pp. 6-7.

En los últimos años Searle ha presentado una distinción entre características intrínsecas y características relativas al observador que puede dar luz acerca de la irreductibilidad de la inteligencia humana al modelo computacional. La atracción de la gravedad, la fotosíntesis o el electromagnetismo son, por ejemplo, materia de las ciencias naturales porque estas ciencias describen rasgos intrínsecos de la realidad. Pero en cambio, una silla, un día estupendo para ir de excursión, un billete de mil pesetas, o la prohibición de fumar en clase no son materias de las ciencias naturales porque no son rasgos intrínsecos de la realidad. Todos esos fenómenos son físicos pero tienen un carácter extrínseco, relativo a un observador que les asigna tal propiedad o los usa de tal manera. Pues bien, mientras las ciencias naturales pretenden "descubrir y caracterizar las realidades que son intrínsecas al mundo natural, (...) por la propia definición de computación y de conocimiento, no hay modo de que la ciencia cognitiva computacional pueda llegar a ser una ciencia natural, porque la computación no es una característica intrínseca del mundo, sino que es relativa a los observadores"¹⁷. Aunque la computación se defina como la manipulación de símbolos, la noción de símbolo no pertenece a la física o a la química, sino que algo es un símbolo sólo porque es usado, tratado o considerado como tal por un observador, que asigna a un objeto físico –un sonido, una inscripción, un gesto– una interpretación. La computación existe sólo relativamente a un agente o un observador que impone una interpretación computacional a un fenómeno determinado.

Me he detenido en esta cuestión pues es la segunda tesis que en estas páginas quiero defender. Hay fenómenos y procesos que consideramos naturales, que aun teniendo una dimensión física, un carácter físico, no tienen una explicación intrínseca, una explicación dentro de sí o en sí mismos. Encontramos la razón de estos fenómenos fuera de sí, extrínsecamente, en su relación con las acciones humanas, con las interpretaciones que de ellos hacemos los seres humanos. El caso de las imágenes del PET es un magnífico ejemplo de estos fenómenos extrínsecos: son los artefactos construidos por la ingeniería médica más avanzada para tratar de comprender mejor el funcionamiento cerebral. Estos fenómenos extrínsecos, que tienen sus razones fuera de sí –en particular, los lenguajes, los símbolos que los seres humanos empleamos– dejan de entenderse si se los aísla de su uso efectivo reduciéndolos a fenómenos meramente intrínsecos, a cosas que pasan. Los lenguajes son artefactos forjados histórica y culturalmente por los seres humanos para facilitar nuestra comunicación y para expresar nuestra intimidad, que llegan a constituir auténticos sistemas de identidad social y que son el humus de nuestra racionalidad. Los argumentos de Putnam y Searle muestran que una explicación reductiva de los fenómenos mentales mediante su asimilación a los procesos computacionales de los ordenadores no explica realmente nada, porque la computación de los ordenadores sólo se explica en su relación con los seres humanos que con ellos interactúan.

17. IDEM, *The Rediscovery of the Mind*, p. 212.

4. *Hacia una comprensión integral de la inteligencia*

Destinaré la última parte a esbozar las líneas por las que discurre una comprensión integral de la inteligencia humana, que está suscitando por doquier en estos últimos años "una cierta euforia"¹⁸. Tras décadas de superespecialización muchas veces tediosa asistimos al regreso del ideal aristotélico de un saber total, en el que se articulen armónicamente los mejores resultados de las diversas ciencias, y en el que, sin borrar las diferencias, se gane en riqueza de perspectivas, en complejidad y, por tanto, en potencia explicativa. En este comienzo de un nuevo siglo, ¿por qué pensar que haya una solución simple a un conjunto de problemas tan complicados como son los relativos a la inteligencia humana y los fenómenos a ella asociados? Por el contrario, parece mucho más certero concebir la verdad, en este campo y en tantos otros, como aquel ideal que finalmente la comunidad científica alcanzará si prosigue sin desánimo su búsqueda cooperativa multidisciplinar.

El funcionalismo, la corriente todavía hegemónica en la filosofía de lo mental que concibe el ser humano como una máquina simbólica, es insuficiente para explicar o para dar razón de cómo pensamos realmente o cómo actuamos inteligentemente los seres humanos, cómo nos enamoramos o por supuesto por qué enfermamos. Por este motivo, resulta de un extraordinario interés el reciente movimiento en la ciencia cognitiva que se aproxima a la biología y a la comunicación social para tratar de comprender mejor la naturaleza de las acciones humanas y del lenguaje, aunque ellos lo hagan con vistas al diseño de una nueva generación de ordenadores que sean realmente adecuados a los propósitos humanos. Un lugar destacado en esta nueva dirección lo ocupa Terry Winograd, quien con diversos colaboradores ha desarrollado un formidable ataque contra la ortodoxia de la inteligencia artificial. Una característica llamativa de su aproximación –denominada "constructivismo hermenéutico"– es su expresa apelación a filósofos como Heidegger, Gadamer, Wittgenstein y Austin¹⁹. Mientras el modelo tradicional de explicar el lenguaje privilegiaba su relación con el conocimiento, Winograd destaca que nuestro conocimiento práctico es más fundamental que la comprensión teórica, que los seres humanos no nos relacionamos con las cosas primariamente teniendo representaciones de ellas. Esto significa privilegiar la dimensión comunicativa, conversacional, pragmática, del lenguaje y del conocimiento, de forma que el criterio relevante no sea la precisión algorítmica de la representación, sino la capacidad del sujeto de relacionarse efectivamente con los demás.

Las computadoras no son máquinas pensantes sino más bien máquinas de lenguaje, máquinas de procesar símbolos de acuerdo con reglas. El lenguaje es pensamiento y –como sostuvo Tomás de Aquino²⁰ y afirma la psicología

18. VIOLETA DEMONTE, "De la complejidad del 'logos'. Y a modo de presentación", *Revista de Occidente* 119 (1991), p. 10.

19. TERRY WINOGRAD Y FERNANDO FLORES, *Understanding Computers and Cognition*. Reading, MA: Addison-Wesley, 1987, p. xii; TERRY WINOGRAD, "Máquinas pensantes: ¿son posibles? ¿lo somos?", *Revista de Occidente* 119 (1991), p. 144.

20. TOMÁS DE AQUINO, *Summa Theologiae* I, q. 78, a. 4.

contemporánea²¹— no hay pensamiento sin imágenes. Pero al mismo tiempo puede afirmarse que no todo lo que consideramos inteligente puede articularse explícitamente de forma simbólica, esto es, que sea posible reducir las destrezas, la intuición y otras formas de conocimiento tácito a hechos y reglas explícitos. Muchas de las críticas más incisivas a la inteligencia artificial se basan precisamente en un análisis fenomenológico del entendimiento humano como "apertura", como "estar disponible" para la acción en el mundo, más que como manipulación de símbolos. De lo que ahora se trata es precisamente de lograr una comprensión más amplia por cuyo medio sea posible recuperar esas dimensiones perdidas, y en el mismo proceso, sea posible entendernos mejor tanto a nosotros como a nuestras máquinas.

De acuerdo con David Chalmers —que es quizá uno de los filósofos más interesantes en este campo— muchos de los problemas de las ciencias cognitivas son fáciles porque pueden ser explicados en términos computacionales o de mecanismos neuronales. Pero el problema difícil (*'the Hard Problem'*) que se resiste a todos los intentos de explicación es el de la consciencia, el de la experiencia consciente, el aspecto subjetivo de toda experiencia que se ha resistido hasta ahora a cualquier explicación²². Algunos creemos que en la comprensión de la articulación creativa de pensamiento y mundo que acontece en el lenguaje, se encuentra la vía de comprensión de ese problema. El rótulo general de esta posición es el de "externalismo": los contenidos de nuestra experiencia, de nuestra consciencia, no dependen de su modelización lógico-matemática ni siquiera de su soporte físico neuronal, sino que sobre todo dependen del medio físico, del entorno que nos rodea.

Como describió bellamente Walker Percy, cuando Hellen Keller —la niña americana sordomuda y ciega— descubre que los toques que hace Ana Sullivan en su mano izquierda son señales, son el nombre del agua de la fuente en la que le introduce el brazo derecho, comienza en ese instante su vida como persona²³. Cuando el niño de escasos meses coge una flor y balbucea mirando hacia su madre "a flo" o algo parecido, en su conducta aún un sonido, una flor y a su madre, siendo él mismo el autor de la unificación de los otros tres elementos. Esta extraña capacidad de aunar, de relacionar elementos tan dispares, es exclusiva del *Homo sapiens*, y es esa exclusividad lo que quizá resulta más incomprensible para muchos científicos.

Los intentos denodados de enseñar el lenguaje de los sordomudos a chimpancés y otros primates superiores muestran con claridad que en su máximo desarrollo la actividad comunicativa de los monos más inteligentes no pasa de los balbuceos pre-lingüísticos del niño de pocos meses reclamando la leche

21. DANIEL DENNETT, "Our Vegetative Soul. The Search of a Reliable Model of the Human Self", *Times Literary Supplement*, 25 agosto 1995, p. 3.

22. DAVID CHALMERS, "Facing Up to the Problem of Consciousness", *Journal of Consciousness Studies* 2 (1995), pp. 201-203; GÜVEN GÜZELDERE, "The Many Faces of Consciousness: A Field Guide", en *The Nature of Consciousness: Philosophical Debates*, pp. 29-30; DAVID CHALMERS, *The Conscious Mind: In Search of a Fundamental Theory*. New York, Oxford University Press, 1996.

23. WALKER PERCY, "La criatura dividida", *Anuario Filosófico* 29 (1996) p. 1146.

materna. La cuestión interesante es: ¿de dónde le sale el lenguaje al niño de dos años que al ver una flor mira a su madre y dice "a flo"? En nuestra cultura pasamos de la neurología a la lingüística, sin explicar ese salto, que incluso en términos evolucionistas resulta tan extraordinario. Los seres humanos aparecen así como unas criaturas divididas entre biología y lingüística sin que se ofrezca una explicación global suficientemente comprensiva del ser humano²⁴. Sin embargo, los estudios del desarrollo cerebral de los bebés humanos muestran con claridad que un cerebro no es un ordenador, que la naturaleza no lo junta todo y luego lo enciende, sino que el cerebro comienza a trabajar mucho antes de que esté terminado y que el proceso de interacción con el medio decide su desarrollo posterior²⁵.

Por todo esto, algunas de las perspectivas más recientes aspiran a comprender el ser humano de forma no reductiva. Lo que el ser humano es se muestra en su cultura, en su interacción comunicativa con los demás, en sus acciones inteligentes. Su subjetividad, su consciencia, su intencionalidad tienen una base biológica pero sobre todo se forjan culturalmente. El reto de la investigación contemporánea es el progreso en la comprensión de cómo se articulan en la vida e inteligencia humanas esas dimensiones computables –aquellas que en la jerga al uso pueden "correr en ordenador"– y sus dimensiones no computables, y al mismo tiempo avanzar en la comprensión de cómo se articulan las dimensiones sociales o sistémicas y aquellas otras dimensiones más personales. La comprensión integrada de los dos ámbitos es en mi opinión la necesidad intelectual más acuciante en nuestra cultura para el siglo XXI. Soy de natural optimista acerca de la capacidad de la razón humana por alcanzar una mejor comprensión de sí misma, del mundo y de las acciones de los seres humanos. Este optimismo tiene un sólido apoyo en la historia general de la humanidad, pero el crecimiento no es inexorable sino que a la postre depende del esfuerzo personal de individuos concretos.

DR. JAIME NUBIOLA
Universidad de Navarra

24. Ibid. p. 1149.

25. J. MADELEINE NASH, "Fertile Minds", *Time*, 10 febrero 1997.