

# JOHN R. SEARLE EN LA TEORÍA DE LA CONCIENCIA

*JAVIER MONSERRAT, Universidad Autónoma de Madrid*

---

Cuando en los libros de psicología cognitiva o neurología se menciona a Searle, esta mención va siempre acompañada del calificativo de “filósofo”, el “filósofo John Searle”. Es, en efecto, un filósofo que comenzó sus trabajos en la filosofía del lenguaje, campo en el que sus obras alcanzaron notoriedad considerable. A fines de la década de los ochenta, sin embargo, Searle comenzó a interesarse por la llamada “teoría de la mente”, o de la conciencia, tal como estaba siendo tratada en el ámbito científico de la psicología y neurología cognitiva, así como de las ciencias de la computación en general. Los científicos de la mente, muy celosos de su condición de científicos, pocas, muy pocas veces, atienden con respeto a las opiniones de un filósofo que se atreva a entrometerse en sus cosas. Pero una de esas pocas veces, una de esas excepciones, es precisamente John Searle (otro filósofo también mencionado con respeto suele ser, por ejemplo, Daniel Dennett). Ciñéndonos al caso de Searle es comprensible que sus opiniones sean consideradas con respeto, ya que, a mi entender, representan siempre un punto de vista serio, riguroso, objetivo, lleno de sentido común, algo así como un espejo luminoso en que se reflejan las diferentes teorías de la mente para quedar valoradas, criticadas positivamente y matizadas, o simplemente reducidas a cenizas por argumentos que es muy difícil ignorar o contrarrestar por contraargumentos apropiados.

Un hito importante en la explicación de la posición propia de Searle en teoría de la mente fue su libro *The rediscovery of the Mind*. Sin embargo, la cátedra más importante para potenciar el eco de sus opiniones ha sido en la última década la colaboración en el *New York Review of Books*. La revisión crítica a que Searle ha sometido en tan prestigioso medio a las obras más relevantes en teoría de la mente ha hecho que su opinión haya llegado a ser imprescindible. Algunas de las revisiones más importantes (en algún caso con las réplicas a que dieron lugar) fueron reunidas por Searle en un volumen titulado *The Mystery of Consciousness*, al que añadió un capítulo introductorio y otro final para exponer su propia opinión (la misma que ya había expuesto en *The rediscovery...*). En este breve comentario vamos a limitarnos a revisar las revisiones presentadas por Searle en este volumen. Nos ayudará a recordar, una vez más, la ponderada opinión del mismo Searle y, al mismo tiempo, a

reflejar sobre ella la posición de autores tan importantes como Francis Crick, Gerald Edelman, Roger Penrose, Daniel Dennett y David Chalmers<sup>1</sup>.

### *La posición de Searle: la conciencia como problema biológico*

El punto de vista de Searle se enmarca en una perspectiva biológica: la conciencia es un fenómeno biológico producido a partir de la ontología propia y la organización estructural-sistémica de los seres vivos. No es que debamos pensar que todos los seres vivos tengan conciencia; pero la conciencia aparece como una propiedad causada terminalmente por ese modo de realidad que llamamos vida, por lo biológico. La conciencia es así una “propiedad emergente” de la organización biológica. “Resumiendo, pues, mi posición general sobre la manera en que la investigación del cerebro puede plantearse la respuesta a las cuestiones que nos intrigan: el cerebro es un órgano como cualquier otro; es una máquina orgánica. La conciencia está causada por procesos neuronales de bajo nivel en el cerebro, y es ella misma un rasgo del cerebro. Puesto que es un rasgo que surge, emergente, a partir de ciertas actividades neuronales, podemos concebirla como una “propiedad emergente” del cerebro. Una propiedad emergente de un sistema es una propiedad que se puede explicar causalmente por la conducta de los elementos del sistema; pero no es una propiedad de ninguno de los elementos individuales, y no se puede explicar simplemente como un agregado de las propiedades de esos elementos” (30).

La posición de Searle es, pues, emergentista. Esto no quiere decir, sin embargo, que una “palabrita” resuelva el enigma de la conciencia. El emergentismo de Searle postula que la resolución al enigma de la conciencia hay que orientarla hacia la investigación acerca de la ontología físico-biológica y la estructura funcional que hace emerger la conciencia desde las redes neuronales. Pero apuntar hacia ese marco de solución como el que en este momento se presenta científicamente más justificado no supone afirmar que ya hayamos dado con la solución. Searle nos habla del misterio de la conciencia porque, en definitiva, decimos nosotros para explicar su pensamiento, no conocemos todavía ni la ontología físico-biológica ni la estructura funcional que hace emerger la conciencia. Si hubiéramos llegado a conocer estas causas de la conciencia podríamos explicar congruentemente tanto la naturaleza de los *qualia* como las propiedades de la actividad psíquica de ellos derivada, tal como podemos describir por fenomenología (percepción, memoria, atención, conocimiento, pensamiento, aprendizaje, lenguaje, emociones e intenciones, etc.).

Desde un enfoque emergentista se piensa, en efecto, que la producción de los *qualia* tiene relación con las redes neuronales, con los engramas

neuronales. Pero la morfología, fisiología y bioquímica de las redes neuronales es tan compleja que distamos todavía mucho de haber descrito la ontología físico-biológica precisa y la estructura funcional derivada, todo ello presente en las redes neurales, que explica la causación real de los *qualia* y de los procesos psíquicos. Searle sabe que el conocimiento superficial del sistema nervioso –digamos en el nivel de una neurología macroscópica que hoy es habitual- no nos ofrece una visión ni de la ontología ni de la estructura funcional de las interacciones neuronales que haga fácilmente explicable el psiquismo. Lo veremos seguidamente.

Searle observa también que debemos tener en cuenta qué hay que entender por conciencia en la ciencia. Negativamente nos dice que conciencia no es un hecho real definido de forma alambicada y compleja por alguno de los autores, por ejemplo, de la filosofía analítica. Conciencia es para la ciencia el hecho real que todos advertimos en nosotros y que puede describirse fenomenológicamente. “No me parece en absoluto difícil dar una definición de sentido común del término: “conciencia” alude a aquellos estados del sentir y del advertir que típicamente dan comienzo cuando despertamos de un sueño sin sueños y continúan hasta que nos dormimos de nuevo, o caemos en un estado comatoso, o nos morimos, o, de uno u otro modo, quedamos inconscientes” (19). Esta definición, en apariencia quizá un poco simplista, debe ser entendida en el sentido de que la ciencia trata de explicar nuestra experiencia fenomenológica, personal e intersubjetiva-social, del hecho de sentir los *qualia* y los procesos psíquicos derivados.

Sin embargo, aparte de la mencionada dificultad básica relativa al conocimiento todavía insuficiente de la ontología y estructura funcional profunda del sistema nervioso, la explicación científica de la conciencia se encuentra hoy con algunas dificultades o problemas añadidos. Searle menciona cuatro problemas.

El *primer problema* es la tradición dualista derivada de la influencia de Descartes y Galileo para quienes una cosa era la ciencia (que versaba sobre el mundo) y otra la realidad mental del alma (inabordable por la ciencia). Para Searle esta tradición está presente todavía en muchos científicos que eluden complicaciones morales, humanísticas, religiosas, al situar el mundo psíquico fuera del mundo físico y fuera del imperio de la ciencia natural. Pero el camino correcto sería para Searle, en cambio, la vía del emergentismo, al buscar la forma en que los procesos cerebrales causan la conciencia. Y aquí surge el *segundo problema*: el que algunas escuelas materialistas consideren erróneamente que hablar de “causación de la emergencia de la conciencia” equivale a admitir alguna forma de dualismo. Muchos, en efecto, no acaban de entender que el emergentismo habla de una emergencia sistémica que no supone la aparición de ningún tipo de entidad real nueva. “Ese estado, nos dice Searle, no es una entidad separada

de mi cerebro; al contrario no es sino un rasgo de mi cerebro en el momento presente” (21). El *tercer problema* mencionado por Searle es el problema de los *qualia*. La conciencia no es sino el conjunto de estados cualitativos que se nos hacen presentes en los *qualia*. En definitiva es el mismo problema que antes hemos mencionado: a saber, el problema de cómo el complejo sistema físico-biológico del S.N, observado objetivamente, puede llegar a producir algo cualitativamente tan específico y extraño como la experiencia subjetiva de los *qualia*, o conjunto de fenómenos de experiencia psíquica subjetiva. La descripción fenomenológica de éstos nos permite constatar la enorme complejidad y matices de la vida psíquica. Y el problema consiste en tratar de encontrar obviamente, como antes decíamos, la ontología físico-biológica y la estructura funcional que puede llegar a producirlos adecuadamente<sup>2</sup>.

El *cuarto problema* es la tendencia actual, muy extendida, a aplicar a la explicación del psiquismo, y de la conciencia, la metáfora computacional de la mente. Searle ha rechazado siempre la idea de que la mente humana sea un computador digital; o, en otras palabras, que la mente sea un programa (software) de ordenador. Es el punto de vista de la inteligencia artificial fuerte, que Searle rechaza (no así la metáfora débil del ordenador). Insiste en que los ordenadores trabajan como sistema manipuladores de símbolos (por ejemplo, sistemas de ceros y unos) en función de unas reglas o programnas (gramática). Los programas son enteramente sintácticos; sin embargo, las mentes tienen contenidos semánticos. Por consiguiente, la pura gramática no es suficiente para justificar los contenidos semánticos. Por esto la mente no es un ordenador, porque tiene algo más, tiene semántica. En consecuencia, el manipulador del símbolos mediante reglas de gramática en el interior de la “habitación china”, contestaría en chino pero no “sabría” chino. Pero Searle nos ofrece aquí una profundización en su argumento clásico de la habitación china. Palabras como electrón o conciencia designan rasgos intrínsecos a la naturaleza; los impulsos eléctricos de un ordenador son también intrínsecos a la naturaleza. Pero las palabras símbolo, sintaxis, gramática o computación no designan fenómenos intrínsecos a la naturaleza: son constructos artificiales referidos siempre a la conciencia del observador que ha construido esos aparatos y sistemas, en su *hardware* y en su *software*. “Este es un argumento, nos dice Searle, distinto del de la habitación china, pero es más profundo. El argumento de la habitación china mostraba que la semántica no es intrínseca a la sintaxis; éste muestra que la sintaxis no es intrínseca a la física” (30).

*Redes neurales en Francis Crick y Gerald Edelman*

Como hemos visto la posición de Searle es biológica: la conciencia emerge del S.N. y de su estructura fundamental, el cerebro. La conciencia, pues, en alguna manera, será causada por las redes neurales; más exactamente por los procesos neurales de bajo nivel, por la ontología más profunda de su estructura físico-neural. Pero lo que las redes neurales producen son los *qualia*, la vida psíquica con todas aquellas propiedades que describimos por fenomenología. El ordenador produce una pura manipulación gramatical de símbolos, sin sensibilidad ni conciencia; pero el cerebro produce algo más: produce los *qualia*. El cerebro manipula activaciones neurales que permiten evocar la palabra “árbol”, pronunciarla emitiendo ciertos fonemas, saber por su contenido semántico qué significa en el mundo real, e incluso evocar la imagen interna que corresponde a la realidad árbol presente en el mundo real. Por ello, el cerebro es mucho más que un ordenador.

Es pues comprensible que Searle sintonice muy especialmente con aquellos autores que presentan una explicación del funcionamiento general de la mente a partir del concepto de redes neurales, que insisten en decir que lo que las redes neurales producen es la vida psíquica como tal, los *qualia*, y que, aunque admiten la utilidad de modelos computacionales concebidos en el marco de la IA débil, rechazan que la mente sea un ordenador en la forma en que lo hace la IA fuerte. Tal es el caso de autores como Crick y Edelman que parecen gozar de todas las simpatías de Searle.

Searle, en contra de algunas interpretaciones incorrectas del libro de Crick, trata de presentar lo que considera su tesis fundamental: “todas nuestras experiencias conscientes se explican por la conducta de neuronas, y son ellas mismas propiedades emergentes del sistema de neuronas” (32). Considera aceptable la forma en que el análisis de Crick conduce a hacernos una idea de lo que sería la “conciencia visual”, pero lamenta que su falta de preparación y consecuente precisión filosófica le lleve a algunos errores y a una cierta confusión en la presentación de su pensamiento<sup>3</sup>. Para Searle es insuficiente la forma en que Crick se refiere al problema de los *qualia* (se refiere a los *qualia* de los demás, sin darse cuenta de que la propia experiencia del *quale* es la esencia del problema mente-cuerpo). Pero el caso es que Crick parece hablar en ocasiones como si quisiera una reducción eliminativa de la conciencia, cuando la dirección que de hecho toma su libro es la de una explicación causal. Esta confusión filosófica le lleva a aceptar argumentos reduccionistas de Paul y Patricia Churchland<sup>4</sup>, cuando él no trata de negar que la conciencia exista, sino de mostrar que la conciencia es una propiedad emergente del cerebro. “Las sensaciones complejas, afirma Searle que dice Crick, son propiedades emergentes que surgen en el cerebro a partir de las interacciones de sus muchas partes” (39).

Crick presenta la gran importancia del “problema de la ligadura” para explicar la conciencia visual. Este problema hace referencia a que la información que compone la imagen (forma, color, profundidad, movimiento...) se encuentra repartida en numerosas localizaciones del cortex visual (y de otras estructuras más antiguas). ¿Cómo ligan todas esas partes de cerebro y se coordinan para producir en un mismo momento la conciencia visual? Crick se hace eco de las hipótesis de Wolf Singer acerca de que una gama de disparos específica, a 40 hercios, sería la responsable de relacionar la actividad al unísono de todas esas neuronas conectadas en una actividad psíquica unitaria. Crick y su colega Christof Koch extenderían algo esa gama de disparos situándola entre 35 y 75 hercios. Searle mira con reservas esta hipótesis, pero insiste en que tampoco en este caso podría decirse que se ha encontrado la explicación final de la conciencia. “Admiro la disposición de Crick a especular, pero la naturaleza de la especulación muestra cuán lejos está aún de la meta. Supóngase que resultara que la conciencia estuviera invariablemente correlacionada con tasas de disparos neuronales de 40 hercios en circuitos neuronales conectores del tálamo y córtex. ¿Sería eso una explicación de la conciencia? No. Así formulada, jamás la aceptaríamos como una explicación. Sin duda la entenderíamos como un imponente paso hacia delante, pero aún seguiríamos preguntándonos: ¿cómo opera, cómo funciona?” (43).

En definitiva, Searle piensa que Crick defiende una tesis emergentista en el marco de una explicación de la conciencia por redes neurales. Una cierta confusión filosófica oscurece a veces su punto de vista, pero es claro que Crick trata de explicar toda nuestra vida consciente con un efecto causal producido por la activación de redes neurales. Crick se mueve en una dimensión explicativa propia de lo que he llamado una neurología macroscópica en que la neurona (y la bioquímica de los neurotransmisores) son sus protagonistas estelares. Pero Searle piensa que este tipo de neurología, aunque bien orientada, dista todavía mucho de haber llegado a explicar qué ontología profunda y qué estructuras funcionales producen los *qualia* y la vida psíquica.

La consideración que Searle hace de Gerald Edelman es entusiasta. “Entre las teorías neurobiológicas de la conciencia que conozco, la más impresionantemente elaborada y la más profunda es la de Gerald Edelman” (46). El enfoque de Edelman es, en efecto, ya mucho antes que Crick, un intento de explicar la naturaleza de la mente a partir de las redes neurales. Su teoría, de naturaleza especulativa, es mucho más completa, pero, en último término, como observará Searle, tampoco ofrece una respuesta a la pregunta crucial, igualmente incontestada por Crick: qué tipo de ontología profunda subneural puede explicarnos que las redes neurales produzcan realmente el mundo de los *qualia* y todos los complejos matices de la vida

psíquica. Sabemos que el psiquismo tiene que ver con las redes neurales; pero todavía no sabemos explicar exactamente por qué<sup>5</sup>.

Edelman concibe una estructura funcional del cerebro que constituiría la base neural que soporta nuestras actividades psíquicas. Su teoría se fundamenta en tres nociones básicas. La primera es la noción de *mapa neuronal*. Está constituido por un conjunto de neuronas que se activan en sistema en dependencia de ciertas topologías receptoras (la piel o la retina). Las neuronas de la retina, al activarse por la actuación del patrón de luz, activan diversos mapas neuronales conectados entre sí que, al activarse en conjunto, serán responsables de la causación de la experiencia psíquica de la imagen (que es un *quale*). A mi modo de entender los mapas y sistemas de mapas, en la concepción de Edelman, se corresponden con lo que yo mismo y otros llaman patrones, engramas, cánones, pautas... neurales que constituyen el soporte de la experiencia psíquica (una imagen, un sonido, una idea o representación cognitiva, una emoción, un recuerdo evocado por la memoria, etc.)<sup>6</sup>. La segunda noción es la de *selección neuronal de grupo*. A medida que el organismo ejerce sus sentidos y actividad cerebral se van seleccionando grupos de neuronas; muchas de las neuronas no son seleccionadas y van desapareciendo. El cerebro funciona siempre por grupos de neuronas. Las neuronas esenciales que forman parte de un mapa o engrama no son fijas; son una neurona de un grupo neuronal. De esta manera, al menos así lo entiendo, los engramas aparecen más flexibles: contienen la activación de todos los grupos neurales que los constituyen pero de una forma más flexible u oscilante, ya que, en ellos, pueden activarse unas u otras neuronas (pensemos que, en efecto, la imagen en diferentes momentos o el recuerdo, o los recuerdos, de una misma imagen presentan siempre una cierta oscilación experiencial). La tercera noción es el *reingreso*. Hace referencia a que los diferentes mapeados, o sistemas de mapas, mandan señales unos a otros por vías paralelas y multidireccionales. Así los mapas presentes en el cerebro están conectados entre sí y presentan precisas estructuras de corrimiento de la activación de unos a otros, en sistemas rigurosos de activación-desactivación y por buses multidireccionales.

Apoyándose en estas tres nociones la teoría de la conciencia de Edelman es una propuesta especulativa acerca de cómo pudo ir construyéndose la mente humana de una forma selectiva y darwiniana. Su sistema es una reconstrucción evolutiva del proceso que, fundándose en la memoria (fijación y reactualización de los mapas, *the remembered present...*), conduce desde la construcción y activación de los sentidos por razones de beneficio adaptativo hasta la formación de las primeras categorizaciones y de las últimas, ya de carácter superior. A través de este desarrollo complejo y relacional de mapeados y categorizaciones se forma lo que Edelman llama la “conciencia primaria” y después la “conciencia de

orden superior”. De esta forma evolutiva el cerebro va configurando poco a poco su “cartografía global” que es una cartografía de engramas.

Como consecuencia de esta activación de los mapas que soportan percepciones en tiempo real, recuerdos de percepciones, emociones, cogniciones referidas a eventos percibidos en el mundo, pensamientos, etc., aparece la compleja actividad psíquica que se evoca en el sujeto a medida que las múltiples vías de reingreso conectan los sistemas de mapas contruidos con una cartografía lógica que permite hacer uso de ellos para sobrevivir en el mundo real.

“¿Qué se puede pensar de todo esto, se pregunta Searle, como hipótesis sobre la conciencia? Como ya va dicho, se trata del intento más acabado y profundo que yo haya visto en la bibliografía neurobiológica sobre el problema de la conciencia. Lo mismo que Crick, Edelman concibe buena parte de su teoría como especulativa, pero tanto mejor. Sin tesis que someter a prueba no hay progreso en nuestro conocimiento. La dificultad capital, no obstante, resulta obvia: hasta ahora, Edelman no ha ofrecido razón alguna para creer que un cerebro que tiene todas esas características, tendría también, y por lo mismo, sentir y conciencia. Recuérdese que todos los rasgos mencionados de la conciencia primaria –categorización perceptiva, valor, memoria, etc.- se suponen entendidos por la mera especificación de su estructura y de las funciones que cumplen. No podemos pensar en ellos como rasgos de conciencia. La idea es que el entero conjunto de sistemas interrelacionados genera conciencia merced a cartografías de reingreso. Pero, por lo descrito hasta ahora, no sería imposible que un cerebro tuviera todos esos rasgos funcionales de comportamiento, incluida la cartografía de reingreso, sin llegar a ser consciente. El problema es el mismo con que nos encontramos antes: cómo pasar de todas esas estructuras y de sus funciones a los estados cualitativos del sentir y el advertir que todos nosotros tenemos, es decir, a lo que algunos filósofos llaman *qualia*” (53-54).

Edelman ha intentado orillar discretamente el problema planteado por la explicación de los *qualia*. Pero Searle, obviamente, no se da por satisfecho. “El problema de explicar los estados cualitativos internos del sentir y del advertir llamados *qualia* no es un aspecto del problema de la conciencia que pudiéramos dejar de lado; es el problema de la conciencia, porque cualquier estado consciente es un estado cualitativo, y *qualia* no es sino un malhadado nombre para la conciencia de todos los estados conscientes” (55).

Edelman ofrece, pues, una teoría de la conciencia fundada en las redes neuronales. Searle no parece dudar de que la aportación de Edelman es la más potente concepción especulativa de cómo esa compleja cartografía de engramas neuronales que se activan y desactivan respondiendo a redes lógicas conectadas por vías de reingreso son la causa de los movimientos

coordinados, del pensamiento congruente, del lenguaje, etc. Searle no parece dudar que esos engramas causen los *qualia* y la vida psíquica. Pero insiste en que Edelman no nos dice exactamente por qué. No dice qué pasa en lo profundo de esas redes neuronales, en su ontología profunda, de tal manera que podamos entender por qué de ellas surge la vida psíquica.

### *Roger Penrose, Kurt Gödel y los citoesqueletos*

El pensamiento de Penrose va, en definitiva, en la misma línea de Crick, de Edelman y del mismo Searle<sup>7</sup>. Esto, al menos, es lo que yo pienso. Sin embargo, Searle realiza una valoración del pensamiento de Penrose en que no parece ver las aportaciones de Penrose de forma tan positiva como sería deseable. A mi modo de entender queda centrado en las diferencias específicas de Penrose, pero sin advertir que lo fundamental es que Penrose piensa, en definitiva, en una línea biologista de carácter emergentista en que la conciencia es un estado sistémico de la ontología profunda físico-neuronal. Al profundizar Penrose en los fenómenos mecanocuánticos que tienen lugar en las redes neuronales no trata de hacer otra cosa que entender el tipo de ontología física de las neuronas que permitiría explicar la emergencia de la conciencia. Bien o mal, Penrose trataría de hacer lo que, a juicio del mismo Searle, no han hecho ni Crick ni Edelman: explicar un poco más, y tentativamente, por qué la red neuronal puede causar la emergencia de algo tan extraño como los *qualia* y por qué posee una ontología profunda que puede hacer explicable la naturaleza “indeterminista” de las funciones psíquicas.

“Todas las obras que estoy abordando, nos dice Searle, tratan de un modo o de otro de la relación entre la conciencia y el cerebro. Crick, Edelman e Israel Rosenfield piensan, como yo mismo, que los procesos cerebrales –probablemente en los niveles de neuronas, sinapsis y grupos neuronales– causan la conciencia” (57). ¿Pero por qué es tan distinto el enfoque de Penrose? “Las discusiones corrientes, la mía incluida, dice Searle, suponen que los computadores se pueden utilizar, al menos, para simular o modelar procesos mentales y suponen también que el tipo de procesos neurobiológicos que describí en el capítulo anterior (la cartografía neuronal al estilo de Edelman) son los fenómenos adecuados para explicar la conciencia. Penrose piensa que si entendiéramos las implicaciones del teorema de Gödel y la importancia de la mecánica cuántica para el estudio de la mente, rechazaríamos ambos supuestos” (58).

Penrose, por tanto, establece dos tipos de tesis: 1) que la mente humana realiza ciertas funciones que no son susceptibles de simulación por ordenador, son las funciones no computacionales de la mente; 2) que la ontología físico-biológico-neuronal que hace posibles estas funciones no computacionales no es sólo el nivel funcional de las neuronas individuales

y sus redes (que sí sería computable), sino un nivel neuronal más profundo donde suceden eventos mecanocuánticos donde regiría la indeterminación cuántica y no serían computables.

Searle y Penrose coinciden en afirmar que la conciencia puede explicarse científicamente, a partir de fundamentos físico-biológicos, y en negar la IA fuerte. Pero difieren en que Searle acepta la IA débil (la mente puede simularse por ordenador, aunque la mente no sea un ordenador) y Penrose niega la IA débil (algunas funciones de la mente ni siquiera pueden ser simuladas por ordenador). Esta negación se fundamenta en ejemplos sacados del pensamiento matemático; el ejemplo fundamental es la presentación del teorema de Gödel, visto desde el punto de vista del problema de la detención, que conduce a Penrose a establecer que las consecuencias a que llega la mente humana en este tipo de razonamientos matemáticos nunca podrían ser alcanzadas por medio de una simulación computacional.

Searle comienza por discutir la versión de la IA débil en Penrose.

La posición de Penrose, como hemos dicho, es defender que existen algunas funciones de la mente humana que no son simulables por un ordenador que funcione computacionalmente. Son las funciones no computacionales de la mente. Searle, en cambio, piensa que las funciones de la mente son siempre simulables por ordenador. La discusión se centra en los ejemplos matemáticos aducidos por Penrose para probar que, en efecto, la mente no es computacional; en concreto la discusión se centra en el teorema de Gödel.

El punto crucial de la discusión radica en observar que la ontología propia del ser humano –dotada de *qualia*, conciencia y actividad psíquica– permite un tipo de operaciones funcionales que en muchos casos serán cuasi-computacionales (recorriendo casi los mismos pasos con que un ordenador aplicaría la secuencia de un algoritmo), pero en otros no. Es lo que sucede, por ejemplo, cuando por la intuición se “ve” la solución de un problema matemático, se alumbra una melodía musical, una vivencia poética o cualquier sentimiento estético. Es evidente que el ordenador no podrá simular nunca el sentimiento de la intuición intelectual que resuelve un problema matemático, o el alumbramiento de una melodía musical o una pauta estética. Pero esto no significa que no pueda programarse a un ordenador para que –de forma mecánica y ciega, aplicando reglas computacionales– pueda simular la forma en que la intuición humana resuelve problemas matemáticos y efectivamente los resuelva. En este punto creemos sinceramente que Searle tiene razón y no vemos razones para admitir que puedan construirse programas de ordenador que simulen integralmente las decisiones y productos del comportamiento humano. Pero, aunque esta simulación se produjera, la ontología y estructuras funcionales del ordenador serían distintas de las de la mente humana. En

otras palabras: la posible amplitud de la simulación de la conducta por ordenador no puede llevarnos a la conclusión injustificada de que el ordenador y el hombre sean lo mismo; a saber, que tengan la misma ontología y que lleguen quizá a idénticos resultados por la misma estructura funcional.

Un ordenador puede simular perfectamente la evolución del tiempo meteorológico, pero a nadie se le ocurre pensar que el ordenador sea el espacio físico real donde las temperaturas, la presión, la velocidad del aire, etc., producirán los sucesos predichos por el programa. De la misma manera, puede existir un programa de ordenador que simule integralmente las decisiones de la mente humana, pero eso no significa que el ordenador sea una mente, ni que haya llegado a esa simulación por los mismos procesos ontológicos y funcionales propios de la mente. Creemos, pues, con Searle, que Penrose se equivoca cuando pretende describir cierto tipo de actividades mentales (en el razonamiento matemático) que no pudieran ser simuladas computacionalmente, aunque ellas mismas, como productos reales de la ontología psíquica de la mente, en efecto no lo sean<sup>8</sup>.

Pero Searle discute también las propuestas cuánticas de Penrose.

Usemos las mismas palabras de Searle para resumir el punto de vista de Penrose. “De acuerdo con Penrose, no podemos descubrir la respuesta al problema de la conciencia en el nivel neuronal, porque las neuronas son demasiado grandes: son objetos explicables en términos de física clásica, y por lo mismo son computables. Penrose cree que, puesto que las neuronas son computables, no pueden explicar la conciencia que no es computable. Tenemos que penetrar en la estructura interna de las neuronas y hallar allí una estructura llamada citoesqueleto, que es el marco que mantiene unida a la célula y constituye el sistema de control de sus operaciones. El citoesqueleto contiene pequeñas estructuras tubiformes llamadas “microtúbulos”, y éstas, según Penrose, desempeñan un papel crucial en el funcionamiento de las sinapsis. He aquí la hipótesis: ‘En la concepción que tentativamente avanzo, la conciencia sería alguna manifestación de ese estado interno citoesquelético, cuánticamente intrincado, y de su implicación en el juego... entre los niveles de actividad clásica y cuántica’. En otras palabras, el material citoesquelético está todo permeado de fenómenos cuánticos, y cuando este micronivel se implica en el macronivel de las neuronas, etc., entonces aparece o emerge la conciencia. Las neuronas no están en el nivel adecuado para explicar la conciencia. Las neuronas podrían ser como mucho “un ingenio magnificador” de la acción real que se da en el nivel citoesquelético” (82).

Pues bien, Searle critica a Penrose porque argumenta la necesidad de recurrir a niveles explicativos cuánticos porque sólo ellos serían no computacionales, siendo así que es criticable y dudoso su rechazo de la IA débil (y que, por tanto, la mente tenga en realidad aspectos funcionales no

computacionales). Critica que pretenda resolver el enigma de la conciencia recurriendo a otros enigmas todavía mayores como la todavía buscada teoría de la gravedad cuántica o una física también por descubrir (por su parte tan místicas o más que la misma conciencia). Según Searle Penrose no aclara, pues, el misterio de la conciencia sino que, al contrario, lo aumenta con otros misterios añadidos. Por último, critica también su interpretación de que el funcionamiento de las neuronas en el nivel macrofísico clásico sea computable y no pudiera nunca por ello dar una explicación cumplida de la actividad propia de la mente (indeterminada y no computable).

Entiendo, como hace Searle y e incluso el mismo Penrose, que la propuesta de explicación cuántica de la conciencia a través de citoesqueletos y microtúbulos no es sino una propuesta tentativa que contiene mucha especulación; es una teoría *in fieri*, es evidente. No obstante mi punto de vista no es tan crítico como el de Searle. Pienso que la referencia a niveles cuánticos no sólo se justifica para dar razón de un pretendido proceso no computacional de la mente (por otra parte discutido); se justifica también para explicar desde una ontología física más apropiada toda la maquinaria funcional del psiquismo que describimos fenomenológicamente (que, por incongruencia con el mecanicismo determinista de la mecánica clásica ha dado lugar a reduccionismos y a monismos). Pienso que iluminar el enigma de la conciencia apuntando a otros enigmas no es pretender resolverlo, sino abrir hipótesis y posibilidades explicativas (una proliferación de teorías en sentido de Feyerabend) que el mismo Searle admite, por ejemplo, al hablar de Edelman, como esencial para la ciencia. Pienso que Penrose al criticar el nivel macroscópico de las neuronas como posible nivel explicativo de la conciencia no hace sino lo que el mismo Searle repite una y otra vez al hablar de Crick y Edelman: que las topografías neuronales no ofrecen una explicación adecuada y satisfactoria de por qué los seres humanos, y los animales, sentimos los *qualia*, la conciencia y la experiencia psíquica global, postulando la necesidad de apuntar a niveles más profundos.

La verdad, apoyamos que no se considere a Penrose como la panacea universal. Es evidente que sus ideas son tentativas. Pero no entendemos esa antipatía que Searle parece mostrar hacia sus propuestas, hasta no dejarle entender que éstas son, hoy por hoy, el camino más serio y prometedor, que haya sido abierto por alguien, para intentar ir más allá de Crick y Edelman, en el misma línea, además, del pensamiento del propio Searle<sup>9</sup>.

### *Inteligencia artificial fuerte en Dennett y Chalmers*

Entendemos que la posición de Dennett es clara, y fuera de toda duda: defiende una interpretación de la mente en el marco de la IA fuerte. El

hombre, la mente, es, pues, un ordenador. ¿Qué papel juega entonces la conciencia? Una interpretación radical de Dennett nos haría responder: la conciencia no existe para la ciencia (en la tradición del objetivismo conductista y fisicalista). Una interpretación más suave sería decir: la conciencia es algo real pero no cumple ningún papel causal en el comportamiento, es algo “epifenoménico”; a lo más, serviría como sistema de control de que todo funciona bien, es decir, de que el ordenador cerebral está realizando cumplidamente sus funciones (por mi conocimiento de autores como Dennett y Kosslyn creo que se podría llegar a pensar que esto último llegan, en último término, a concederlo para salvaguardar una cierta funcionalidad evolutiva de la conciencia).

La revisión de Dennett hecha por Searle se refiere a *Consciousness Explained*, la obra en que parece llegar a su doctrina final sobre la conciencia<sup>10</sup>. La virtualidad del análisis de Searle consiste en que le hace decir a Dennett con toda claridad y dureza lo que Dennett dice rodeado de un innumerable sistemas de amortiguadores, una enorme cantidad de tinta de camuflaje con en el calamar y, en definitiva, mareando la perdiz todo lo que puede. Pero al resplandecer con toda crudeza la tesis de Dennett se intuyen también en toda su fuerza las insuficiencias, e incluso la falta de “sentido común”, que en ellas se implican.

Para Searle la ciencia debe explicar los datos empíricos: y el primer dato empírico es la existencia de la conciencia y de los estados subjetivos internos. Pues bien, Dennett comienza negando la existencia de los datos. “El propósito principal del libro de Dennett es negar la existencia de estados mentales internos y ofrecer una concepción alternativa de la conciencia, o mejor, de lo que él llama ‘conciencia’” (96). “Si Dennett niega la existencia de estados conscientes, según los concebimos de ordinario, ¿cuál es su tesis alternativa? No es para nada sorprendente que sea una versión de la IA fuerte. A fin de exponer esa versión tengo que explicar primero, nos dice Searle, cuatro nociones de las que él se sirve: las máquinas de Von Neumann, el conexionismo, las máquinas virtuales y los memes” (97).

La crítica de Searle a Dennett es muy fuerte. Dennett presenta una pretendida explicación científica que, por no ser congruente con los hechos, acaba por negarlos. El de Dennett es un discurso científico absolutamente injustificable si nos atenemos a las prescripciones epistemológicas para construir con rigor una ciencia que nos explique el mundo real.

Veamos cómo nos resume Searle su valoración crítica.

“El asunto es el siguiente. La ciencia aspira, ciertamente, a la objetividad epistémica. Se propone conseguir un conjunto de verdades que estén libres de nuestros particulares prejuicios y preferencias. Pero la objetividad epistémica del *método* no exige la objetividad ontológica del *objeto de estudio*. Es un hecho objetivo –en sentido epistémico- que yo y

otros como yo tienen dolores. Pero el modo de existencia de esos dolores es subjetivo –en el sentido ontológico-. Dennett tiene una definición de ciencia que excluye la posibilidad de que la ciencia pueda investigar la subjetividad, y cree que la objetividad de tercera persona de la ciencia le obliga a esa definición. Pero esto no es sino un mal juego de palabras con “objetividad”. El fin de la ciencia es conseguir una explicación sistemática de cómo funciona el mundo. Una parte del mundo consiste en fenómenos ontológicamente subjetivos. Si tenemos una definición de ciencia que nos prohíbe investigar esa parte del mundo, lo que ha de cambiarse es la definición, no el mundo” (108). Si nos atenemos a la interpretación de Searle la posición de Dennett sería, pues, fisicalista-conductista y, además, reduccionista al explicar el comportamiento humano por puro *software* computacional, dejando la conciencia (y la actividad psíquica consciente) reducida a un puro epifenomenalismo<sup>11</sup>.

Pero Dennett no es el único defensor de la IA fuerte.

Chalmers defiende también una enrevesada interpretación de la conciencia que cabría también en el marco de la IA fuerte con influencias de una también enrevesada interpretación del funcionalismo<sup>12</sup>. Cuando uno lee el libro de Chalmers, y se toma nota de sus propuestas y conclusiones, no queda otra alternativa que la perplejidad. Pero la perplejidad es mayor cuando se ve que el libro de Chalmers ha sido leído y ponderado, hasta el punto de hacerse un punto de referencia que se hace visible, por ejemplo, en el hecho de que sea uno de los editores del volumen de trabajos de la tercera conferencia de Tucson, Arizona, sobre la conciencia<sup>13</sup>.

Searle le somete a una crítica que alguno calificaría de “despiadada”, si es que puede considerarse así a una crítica que consiste simplemente en decir con claridad, y sin marear la perdiz, lo que el libro realmente dice. Y lo que en realidad hace Chalmers es ofrecer un concepto de conciencia como “sistema de información” que conduce a pensar, si nos lo tomamos en serio, que un termostato, un circuito de calefacción o refrigeración, o un mismo coche, son sistemas a los que podemos atribuir propiamente la “conciencia”. Pero para llegar a conclusiones tan absurdas se presenta un complejo recorrido por la IA fuerte, por el “dualismo de propiedades” y por la teoría funcionalista de la mente que constituye una enrevesada presentación del menú capaz de satisfacer la frivolidad intelectual de muchos de los asiduos a la teoría de la mente.

“Algunos libros, nos dice Searle, son importantes no porque resuelvan algún problema, o porque lo planteen de un modo que permita vislumbrar una solución, sino porque son sintomáticos de la confusión de una época. El libro de Chalmers ha sido saludado como un gran paso adelante en la filosofía de la mente; fue muy debatido entre centenares de académicos que acudieron recientemente a un simposio sobre la conciencia en Tucson; ha

sido citado en las páginas del semanario *Time*; y la faja del libro reproduce encomios de varios filósofos famosos” (149).

### *Cómo hacer un “problema” del “misterio” de la conciencia*

En el último capítulo del libro que comentamos vuelve Searle sobre su posición personal ante el enigma de la conciencia. Podríamos resumirla, haciendo un esfuerzo por expresar con claridad y orden lo que se encuentra más disperso en el libro de Searle, en algunos puntos que presentan una gran coherencia lógica. Personalmente nos sentimos muy identificados con las posiciones de Searle que, como antes decíamos, son algo así como el “sentido común inglés” aplicado a la teoría de la mente<sup>14</sup>.

1) Es un hecho empírico de experiencia subjetiva, personal, pero también consensuado socialmente, que advertimos en nosotros la presencia de *qualia*, conciencia y, en general, la actividad psíquica de conjunto. Este hecho debe ser explicado por la ciencia y no tiene justificación epistemológica ignorarlo.

2) Este hecho juega un papel esencial en el funcionamiento de la mente humana. Aunque es verdad que la mente tiene niveles funcionales inconscientes muy amplios, la experiencia nos dice que tener conciencia y actividad psíquica subjetivo-conciente representa un modo de ser ontológico y funcional que impone a la mente humana propiedades esenciales de su funcionamiento. La ontología y estructura funcional de un ordenador, por ejemplo, es distinta a la humana porque no posee psiquismo (no posee ontología o *hardware* psíquico) y no puede contar con él para implementar su funcionamiento (no posee *software* psíquico).

3) La primera pista científica para investigar las causas y la naturaleza del psiquismo (de la conciencia y de los *qualia*) es la biología que, a su vez, conecta con la física. Esta vía de investigación es la que, con el tiempo, permitirá pasar del enigma de la conciencia al problema de la conciencia. Pasar a un problema progresivamente clarificado. Autores como Crick, Edelman, Rosenfield y el mismo Penrose (y otros muchos que Searle no menciona ni hemos comentado, por tanto, en este artículo) van en esta línea, con mayor o menor acierto.

4) Searle considera que lo que hasta ahora sabemos por la neurología permite situar los niveles explicativos que causan la actividad psíquica en redes neuronales (engramas, cánones, patrones, pautas, etc.) que se construyen en densas estructuras lógicas de activación-desactivación (que responden a las cartografías de Edelman) que, quizá (Searle no es entusiasta de las propuestas de Penrose), escondan en su ontología más profunda fenómenos cuánticos intrínsecamente relacionados con la conciencia y sus funciones.

5) La IA débil puede ser de mucha ayuda para entender tanto el funcionamiento neurológico de esas redes neuronales como la estructura de las funciones psíquicas (y sus mecanismos inconscientes subyacentes). La IA débil puede simular, por ejemplo, las formas de conexión-desconexión (activación-desactivación) de las redes neuronales y sus módulos específicos. Edelman y su equipo han construido varios robots de simulación de redes neurales. Pero esto no significa que se haya podido simular hasta el momento los efectos psíquicos, los *qualia*, la conciencia, que en los seres vivos reales producen las redes neuronales. Las estructuras lógicas de la actividad psíquica pueden también ser simuladas; un robot, por ejemplo, puede conversar con nosotros simulando las contestaciones que nos daría otro ser vivo de nuestra especie. Searle piensa, en contra de Penrose, que no existen de principio las llamadas actividades no computacionales de la mente que no pudieran ser en absoluto simuladas. Sin embargo, insistimos, la simulación también sería en estos casos realizada por mecanismos computacionales ciegos, sin conciencia, merced a los programas contruidos por un ser vivo con conciencia a partir de la experiencia de la actividad psíquica.

6) ¿Podría en absoluto darse un ordenador con conciencia, es decir, con una ontología semejante a la del ser vivo? Lo que sabemos con certeza es que los computadores actuales no tienen una ontología con conciencia y, por tanto, diseñan sus programas de procesamiento adaptados a ese tipo de ontología mecánica y ciega (el ser vivo, en cambio, diseña evolutivamente otros programas porque cuenta con una ontología consciente). Sin embargo, no existe ninguna dificultad teórica para Searle en concebir que en el futuro se pudiera construir un ordenador con un tipo de ontología (*hardware*) que causara la conciencia y que contara con ella en el diseño de sus programas funcionales. En todo caso es ahora difícil pensar en esa posibilidad cuando todavía no conocemos bien cómo la realidad biológica produce la conciencia real de los seres vivos.

7) Estas consideraciones llevan a la conclusión inequívoca de que la IA fuerte, en el sentido de Dennett y Chalmers (y otros muchos autores no mencionados) no puede ser aceptada como propuesta epistemológica y científicamente viable. Los ordenadores actualmente existentes, bien sean seriales (Searle les llama máquinas de Von Neumann) o conexionistas, no tienen, primero, una ontología semejante al ser humano y, segundo, funcionan con unos programas adaptados a su ontología específica y, por ende, distinta de la del ser humano. Sin embargo, tal como antes decíamos, para simular tanto las redes neurales como el funcionamiento psíquico la IA débil puede proporcionar modelos útiles que ayuden a entender el funcionamiento específico del psiquismo. Pero dejando siempre establecido el principio de que no existen bases científicas para pensar que el hombre sea un ordenador (identidad de *hardware*) y ni para pensar que sus

funciones psíquicas sean también semejantes al programa de un ordenador (identidad de *software*).

No quiero concluir sin mencionar unas palabras de Searle, sin duda muy interesantes, en relación a la numerosa correspondencia mantenida con ocasión de sus artículos en el *New York Review of Books* y al extraordinario apasionamiento, a mi modo de ver pseudocientífico, de que hacen gala muchos defensores actuales de la IA fuerte.

“Curiosamente, he hallado más pasión en los partidarios de la teoría computacional de la mente que en los fieles de las doctrinas religiosas tradicionales sobre el alma. Algunos computacionalistas dan una intensidad casi religiosa a su fe en que nuestros problemas más profundos sobre la mente tendrán una solución computacional. Aparentemente, mucha gente cree que, de uno u otro modo, a menos que se demuestre que somos computadores, perderemos algo terriblemente importante. No estoy seguro de entender la fuente de la intensidad de esos sentimientos. Roger Penrose observa también que cuando trató de refutar la hipótesis computacional de la mente, sus argumentos fueron recibidos por alaridos de indignación. Yo sospecho que estos sentimientos tan fuertes podrían proceder de la convicción, compartida por mucha gente, de que los ordenadores son la base de un nuevo tipo de civilización: de una nueva manera de dar sentido a nuestras vidas, de una forma nueva de entendernos a nosotros mismos. El computador parece suministrar al menos un modo de explicarnos a nosotros mismos conforme a la imagen científica del mundo. Y lo que tal vez sea más importante, la teoría computacional de la mente expresa una cierta voluntad tecnológica de poder. Si podemos crear mente con sólo diseñar programas de computación, habremos consumando el dominio tecnológico final del hombre sobre la naturaleza” (169-70).

Quisiera añadir algo a esta explicación de Searle en torno al apasionamiento de los computacionalistas. Para explicar lo que pasa creo que no debemos olvidar lo que hoy constituye la enorme especialización de los grandes profesionales de la investigación universitaria y científica. Hay quienes se han especializado en neurología, pero hay también quienes se han especializado en matemáticas, en computación, en modelos matemáticos aplicados a las CCHH, en visión artificial, en teorema Fourier, en inteligencia artificial, en ingeniería del conocimiento, en modelos expertos, etc. El poder y prestigio del “psicólogo matematizado” tiende obviamente a ampliar su área de dominio, es decir, el área de dominio que sólo él es capaz de controlar porque sólo él domina las técnicas que permiten hacerlo. De que el hombre sea o no sea un computador depende un crucial balance del prestigio y poder intelectual; en definitiva, depende incluso el interés del mercado por las ideas que uno mismo está produciendo.

En definitiva, en teoría de la mente hay hoy dos grandes orientaciones que disputan entre sí. Primero están los que explican la mente por redes neuronales y se fundan en los resultados empíricos de la neurología. Segundo están quienes explican la mente como puros programas de ordenador semejantes a los actuales. Searle, Crick, Edelman, Penrose, otros muchos, entre ellos, modestamente, un servidor de ustedes y, lo que es más importante, el “sentido común científico” parece que están con los primeros. Pero... ¡cuidado con los segundos! Son un poco respondones y, si no, vean el caso de Dennett.

Javier Monserrat  
C/ Universidad Comillas, 7  
28049 MADRID

---

## NOTAS

- 1) Cfer. Searle, John R., *The Rediscovery of Mind*, MIT, Cambridge Massachusetts 1992; Searle, *The Mystery of Consciousness*, Granta Books, London 1997; en español: *El misterio de la conciencia*, Paidós, Barcelona 2000. Al citar textos de Searle daremos entre paréntesis las páginas de esta edición en español.
- 2) La posición teórica de Searle es emergentista. Sin embargo, no es fácil delimitar conceptualmente el ámbito explicativo del emergentismo. He intentado aportar algo a esta delimitación en: Monserrat, J., *Epistemología evolutiva y teoría de la ciencia*, P. Universidad Comillas, Madrid 1987, cap. VI; Monserrat, J., *La Percepción Visual. La arquitectura del psiquismo desde el enfoque de la percepción visual*, Biblioteca Nueva, Madrid 1998, cap. XIV. En estos libros coincido con Searle en observar la tendencia del Identismo a reducir erróneamente el emergentismo a dualismo, sin advertir que el emergentismo no admite, como dice Searle, “entidad o actividad separada”.
- 3) Cfer. Crick, Francis, *The Astonishing Hypothesis: The Scientific Search for the Soul*, Simon and Schuster, New York 1994; en español: *La búsqueda científica del alma*, Debate, Madrid 1994. He tratado más ampliamente la posición de Crick en: Monserrat, J., *Francis Crick y la emergencia de la conciencia visual*, en: *Pensamiento* 52 (1996) 241-252, donde coincido con Searle en entender como emergentista la posición de Crick, así como en poner de manifiesto los aspectos filosóficos débiles de su análisis.
- 4) Cfer. Churchland, Patricia S., *Neurophilosophy. Toward a Unified Science of the Mind/Brain*, MIT, Cambridge Massachusetts 1989; Churchland, P.S., Sejnowski, T.J., *The Computational Brain*, MIT, Cambridge Massachusetts 1992.
- 5) Para seguir el pensamiento de Edelman en relación a la mente hay tres libros fundamentales: Edelman, G., *Neural Darwinism: The Theory of Neuronal Group Selection*, Basic Books, New York 1987; Edelman, *The Remembered Present: A Biological Theory of Consciousness*, Basic Books, New York 1989;

- Edelman, *Bright Air, Brilliant Fire: On the Matter of the Mind*, Basic Books, New York 1992.
- 6) Véase mi artículo: Monserrat, J., *Engramas neuronales y teoría de la mente*, en: *Pensamiento*, vol. 57, n. 218 (2001) 177-211.
  - 7) Cfer. Penrose, Roger, *La nueva mente del emperador*, Mondadori, Barcelona 1991; Penrose, *Las sombras de la muerte. Hacia una comprensión científica de la conciencia*, Crítica-Drakontos, Barcelona 1994.
  - 8) Al igual que Searle, también nosotros hemos criticado la pretensión de Penrose en su rechazo de la IA débil: Monserrat, J., *Penrose y la mente computacional*, en: *Pensamiento* vol. 55, n. 212 (1999) 177-216.
  - 9) He presentado una consideración de las propuestas cuánticas de Penrose más positiva que la de Searle en: Monserrat, J., *Penrose y el enigma cuántico de la conciencia*, en: *Pensamiento*, vol. 56, n. 215 (2000) 177-208.
  - 10) Cfer. Dennett, D., *Consciousness Explained*, Little, Brown and Company, Boston 1991. Hay traducción en español.
  - 11) La réplica de Dennett a Searle, incluida en el libro (así como la contraréplica de Searle) es violenta e intemperante. Dice cosas tan sorprendentes como atribuir a Edelman una posición de IA fuerte, siendo así que Edelman, como hemos comprobado, repite una y otra vez, hasta la saciedad, que la lógica funcional de los mapas neuronales no es la del ordenador. El hombre no es para Edelman un ordenador, aunque la lógica funcional de las cartografías neuronales pueda simularse en diseños (o robots) de IA débil, tal como el mismo Edelman con su equipo han intentado en los programas Darwin I, II y III.
  - 12) Cfer. Chalmers, David J., *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press, 1996; en español: *La mente consciente*, Gedisa, Barcelona 1999.
  - 13) Cfer. Hameroff, J.R., Raszniak, A.W., Chalmers, D.J., *Toward a Science of Consciousness III. The Third Tucson Discussions and Debates*, MIT, Cambridge Massachusetts 1999.
  - 14) Dennet cuenta en uno de sus libros cómo Gilbert Ryle le contó que en uno de sus viajes coincidió en el tren con Bertrand Russell y éste le dijo que el “sentido común” era algo que había inventado John Locke y que, desde entonces, sólo lo habían tenido los ingleses (Dennett, D., *La peligrosa idea de Darwin*, Galaxia Gutemberg, Barcelona 1999, pags. 30-31). Recordemos que Searle es inglés aunque trabaje como profesor en Berkeley.