

# El cerebro, el yo y el libre albedrío: Una discusión entre Llinás, Munévar y Searle\*

*The brain, the self, and free will:  
A discussion between Llinás, Munévar and Searle*

JOSÉ ANDRÉS FORERO-MORA\*\*

Corporación Universitaria Minuto de Dios. jforero@uniminuto.edu

LAURA I. GIRALDO CEBALLOS\*\*\*

Corporación Universitaria Minuto de Dios. laura.giraldo@uniminuto.edu

RECIBIDO EL 10 DE FEBRERO DE 2021, APROBADO EL 16 DE AGOSTO 2021

## RESUMEN      ABSTRACT

El presente texto pretende revisar críticamente las relaciones entre el cerebro, el yo y el libre albedrío a partir de tres propuestas teóricas. En primer lugar, se analiza la posición de Llinás, según la cual el yo, y con ello nuestro libre albedrío, es un mito. Como respuesta a esta posición negativa, se analiza en segundo lugar la posición de Munévar, según la cual el yo debe ser entendido como el cerebro en su complejidad y, en tercer lugar, la propuesta naturalista de Searle, según la cual el yo no debe entenderse como la actividad cerebral en su conjunto, sino como una propiedad sistémica que emerge de él. En ambos casos se evalúan las consecuencias que estas tienen posiciones sobre la relación entre el cerebro y el yo para el libre albedrío.

This text aims at critically reviewing the relationships between the brain, the self and free will based on three theoretical proposals. In the first place, the position of Llinás according to which the self, and with it the free will, is a myth is analyzed. In response to this negative position, the position of Munévar, according to which the self should be understood as the brain in its complexity, is analyzed in the second place. In the third place, the naturalistic proposal of Searle according to which the self should not be understood as the activity of the brain as a whole but as a systemic property that emerges from it is analyzed. In both cases, the consequences that these positions have on the relationship between the brain and the self for free will are evaluated.

## PALABRAS CLAVE      KEY WORDS

Libre albedrío, conciencia, yo, emergentismo.

Free will, consciousness, self, emergentism.

\*Agradecemos al profesor Carlos Patarroyo por la lectura de una versión previa de este artículo y sus útiles sugerencias.

\*\*  [orcid.org/0000-0003-1940-4024](https://orcid.org/0000-0003-1940-4024) **Google Scholar**

\*\*\*  [orcid.org/0000-0002-0715-1835](https://orcid.org/0000-0002-0715-1835) **Google Scholar**



El libre albedrío es quizá uno de los problemas más antiguos de la filosofía. En la antigüedad estaba directamente relacionado con la noción de destino (Estoicos): si estamos incluidos en un plan general del destino, nuestras acciones futuras ya están determinadas, pero, si esto es así, ¿realmente somos libres? Para los medievales, y algunos modernos, el mencionado problema se relaciona, en cambio, con la noción de *omnisciencia divina*: si Dios es omnisciente y conoce a la perfección lo que haremos en el futuro, entonces nuestras acciones futuras están determinadas<sup>1</sup>. La versión contemporánea del problema ha sido enunciada, entre otros, por John Searle: el problema del libre albedrío surge básicamente porque tenemos dos fuertes intuiciones encontradas que consideramos verdaderas: por un lado pensamos que las explicaciones de los fenómenos naturales han de ser completamente deterministas y, al tiempo, estamos convencidos de que actuamos libre y voluntariamente, es decir, que nuestras acciones no están *suficientemente* determinadas por causas externas a nosotros (*Rationality* 270; *Libertad* 27-28; *La mente* 271-272)<sup>2</sup>. El problema del libre albedrío aparece cuando nos consideramos como seres dentro del mundo, cuando, teniendo en cuenta que no podemos renunciar a la verdad de ninguna de las dos intuiciones, las enfrentamos.

De acuerdo con Robert Kane, nuestro sentimiento del libre albedrío descansa en el hecho de que nos vemos como agentes capaces de influenciar el mundo de diferentes maneras. Sentimos que se presentan ante nosotros cursos alternativos de acción, entre los cuales deliberamos y escogemos uno. “Creemos que tenemos libre albedrío cuando nos vemos como agentes capaces de influenciar el mundo de varias maneras” (*A Contemporary* 6).

En principio, si el determinismo es verdadero, entonces no actuamos libremente. Del mismo modo, si la primera intuición es falsa, si el

---

<sup>1</sup> En esta formulación hace falta un paso que, aunque siempre se da por sentado, no sobra hacerlo explícito: Dios por su suma perfección no puede estar equivocado. Si actuáramos de una manera diferente a como Dios ‘sabe’ de antemano que lo haremos ello haría que el conocimiento de Dios fuera falso, es decir, que Él estuviera equivocado, lo cual sería contradictorio con su propia esencia.

<sup>2</sup> Robert Chisholm también presenta una descripción clara y completa del problema tal y como ha sido tomado contemporáneamente, pero en ella involucra el problema de la responsabilidad moral: “Los seres humanos son agentes responsables; pero este hecho aparece en conflicto con una visión determinista de la acción humana y también aparece en conflicto con una indeterminista para la acción” (24). En el presente trabajo nos centraremos en el problema del libre albedrío, no nos ocuparemos del problema de la responsabilidad moral. No obstante, muchas de las conclusiones que obtenemos pueden ser extendidas al problema de la responsabilidad moral.

indeterminismo es verdadero, tampoco podríamos decir que tenemos libertad de la voluntad pues todo estaría indeterminado y estaría sujeto al azar, incluso nuestras acciones. Coincidimos con Gary Watson en que

Lo que destruye la libertad es la falta de autodeterminación, que resulta cuando la voluntad es determinada por otros eventos o estados de cosas y cuando no está del todo determinada. El requerimiento negativo de que la voluntad no esté causalmente determinada por eventos antecedentes está dictado por el requerimiento positivo de que la voluntad sea determinada por sí misma. ("Free Will" 178)

El problema del libre albedrío exige un yo, un 'sí mismo' que determine las acciones. Sin embargo, esto no es condición suficiente para la existencia de la libertad, es necesario que dicho yo tenga acceso a cursos alternativos de acción, a posibilidades alternativas y que pueda elegir entre ellas.

Con respecto al yo, también existen posiciones encontradas. Por un lado, hay estudiosos que afirman que las investigaciones neurobiológicas muestran que el yo es una ilusión o un mito. Por el otro, hay quienes sostienen, con base en estas mismas investigaciones, que es posible defender que hay yo que actúa libremente. Aquí se abren dos posibilidades, por un lado, identificar el cerebro y el yo, aceptando que es el cerebro como sí mismo el que toma decisiones y es causa de las acciones; por el otro, sostener que el yo emerge de los procesos cerebrales complejos, pero que es irreducible a estos. Este texto se sitúa en esta controversia, mostrando que es posible abordar la relación entre el yo y el libre albedrío desde una aproximación neurobiológica. Para esto, ponemos en discusión las propuestas teóricas de Rodolfo Llinás (sección 1), Gonzalo Munévar (sección 2) y John Searle (sección 3) quienes sostienen respectivamente cada una de estas posiciones.

### **Llinás: el mito del yo**

En el capítulo 6 de su libro *El cerebro y el mito del yo*, Llinás responde directamente a la pregunta: qué es el 'sí mismo'. El neurocientífico colombiano plantea los siguientes dos ejemplos que ilustran su posición:

El primero es el concepto del 'Tío Sam'. Al leer en el periódico 'el Tío Sam bombardeó Belgrado' se comprende que las fuerzas armadas de los Estados Unidos se han desplegado hacia esa ciudad, aunque ninguna entidad

corresponda al Tío Sam . . . El segundo ejemplo referente a la afición deportiva resulta aún más interesante. Consideremos los disturbios europeos o sudamericanos asociados con los partidos de fútbol. Es interesante que para los aficionados fanáticos su equipo representa una extensión de ellos mismos, a tal punto que lucharán y arriesgarán su integridad personal por defender 'su equipo', como otros podrían hacerlo por defender su vida, la de sus seres queridos o quizás sus ideales o su fe. (El cerebro 149-150)

Para Llinás es claro que los términos en cuestión 'Tío Sam' y 'mi equipo de fútbol' no refieren a una entidad concreta del mundo. Hay, sin embargo, una diferencia entre estas dos expresiones que Llinás pareciera no ver y que es preciso resaltar. En el segundo caso, el de mi equipo de fútbol, sí hay efectivamente una entidad: hay un número concreto de jugadores que actúan y que son discernibles. La cuestión está en el *posesivo*, pues esa entidad que constituye el equipo de fútbol *N* es mía. En cambio, al leer en el periódico 'el tío Sam bombardeó Belgrado,' no es para nada fácil discernir quién o qué ha actuado allí. No es solo el piloto que suelta la bomba, pues este no lo hace a menos que reciba una orden, que no se da a menos que el presidente la apruebe, y él no lo aprueba a menos que el congreso lo haga, etc. En este sentido, la expresión 'Tío Sam' no refiere una entidad concreta en el mundo. Este mismo rasgo, según Llinás, tiene el yo. "El yo, aquello por lo que trabajamos y sufrimos, es tan solo un término 'útil' referente a un evento tan abstracto como lo es el concepto del Tío Sam respecto de la realidad de algo tan complejo y heterogéneo como son los Estados Unidos" (Llinás, *El cerebro* 149). Para Llinás, el yo no es una entidad concreta en el mundo, sino que es un término útil con el que nos referimos a un evento abstracto, que nos sirve para hablar de una actividad tan heterogénea y compleja como la cerebral.

Ahora bien, es claro que las cualidades secundarias de los sentidos (los olores, sabores, etc.) son invenciones, construcciones o abstracciones de la semántica del sistema nervioso central (SNC). Así por ejemplo, el cuaderno que estoy viendo en este momento sobre mi escritorio no es estrictamente rojo, sino que cierta disposición de las partículas, así como la luz que se refleja, irradia ciertas frecuencias que mi cerebro percibe como rojo. El color rojo como tal no existe, es una invención (construcción o abstracción) de nuestro SNC. Otros animales con capacidad cerebral de identificar más o menos frecuencias de refracción de luz seguramente no verán el cuaderno como rojo. Según Llinás:

La abstracción conocida como el 'sí mismo' no se diferencia fundamentalmente de las cualidades secundarias de los sentidos; el 'sí mismo' es una invención del SNC. Existe dentro del sistema como un polo de atracción, un remolino cuya única existencia real es la que le imparte el ímpetu de múltiples partes dispersas. Es un organizador de percepciones derivadas intrínseca y extrínsecamente: es también el telar en el que se teje la relación entre el organismo y la representación interna del mundo externo. (El cerebro 150)

La razón por la cual Llinás considera al yo como un mito es porque no corresponde a una entidad concreta del mundo, compuesta de partículas físicas; sino que es una abstracción de nuestro SNC para referirse al conjunto de procesos que ocurren en nuestro cerebro. No hay en el cerebro una neurona ni un mecanismo específico que se encargue de realizar las funciones que normalmente le atribuimos al yo, no existe una estructura cerebral específica que le corresponda<sup>3</sup>. En un momento dado son determinadas neuronas y estructuras cerebrales las que interactúan unas con otras, pero un instante después son otras estructuras y otras neuronas las que se encuentran involucradas.

Pero, ¿qué tan justificado está el paso entre los hallazgos neurobiológicos, que permiten decir que no hay una estructura fija que cumpla las funciones del 'sí mismo', y la afirmación de que este es un mito? Uno de los filósofos que ha cuestionado fuertemente este paso es Gonzalo Munévar. Su crítica más aguda está dirigida básicamente al uso de abstracción que hace Llinás. Para este filósofo parece sensato afirmar que las cualidades secundarias de los sentidos constituyen una abstracción, siempre y cuando tengamos en cuenta que esta es una forma corta de decir que la percepción incluye la función de abstraer. Pero de ahí no se sigue que los objetos de percepción sean *objetos* abstractos (Munévar, La evolución 265). Si bien es cierto que el rojo constituye una abstracción de nuestro SNC, sería complicado sostener que las partículas y la refracción de luz que generan mi percepción de rojo sean objetos abstractos. En suma, como analizaremos y discutiremos más adelante, la crítica de Munévar sugiere hacer una distinción entre la percepción que tenemos del yo y el yo: el hecho de que nuestra percepción del yo sea una abstracción no nos autoriza a inferir que el yo sea un objeto abstracto. Por ahora, hemos de analizar qué tan acertada es la crítica que propone Munévar hacia el negativismo de Llinás.

<sup>3</sup> A este respecto es interesante ver el ataque que emprende Llinás contra los defensores de las 'llamadas células de abuela' y de los mapas espaciales del cerebro punto a punto (Llinás, *El cerebro*: 130-140).

Llinás podría defenderse afirmando que el yo es una experiencia y no una entidad. En esto se parece a los colores, ya que estos no son entidades en sí mismas (no existe ahí afuera en el mundo 'el rojo') sino que son experiencias. Munévar afirma que hay que distinguir la experiencia que se tiene y el objeto experimentado, este segundo sí constituye un objeto real. Para ambos pensadores es cierto que no hay una célula organizadora o central (ni un conjunto estable de ellas) que constituya el yo, pero para Munévar esto no implica que el yo no sea una abstracción, ya que hay efectivamente un conjunto de mecanismos que hacen posible esta experiencia, así estos no siempre sean los mismos. No hay una estructura cerebral fija que constituya el 'sí mismo'; por el contrario, existen relaciones dinámicas entre distintos grupos de neuronas y estructuras cerebrales que cumplen con las funciones que normalmente le hemos atribuido al yo. Al decir que un concepto, por ejemplo 'perro', es una abstracción, no estamos negando que existan perros particulares, y, al mismo tiempo, el que haya perros particulares no quiere decir que el concepto 'perro' no sea efectivamente una abstracción. En este sentido, para escapar de la crítica de Munévar, Llinás se vería obligado a aceptar que, en alguna medida, el yo existe, esto es, que no se trata de una mera abstracción o que no es equivalente a afirmar que es algo así como un objeto no existente.

### **Munévar: el cerebro como sí mismo**

De acuerdo con Munévar, la noción de yo no hace referencia solamente a los procesos conscientes del cerebro, sino también a los procesos inconscientes. Según él, esta noción más amplia del yo permite dar cuenta de una manera clara de la autodeterminación de las acciones y con ella del libre albedrío. No obstante, como observaremos, esta noción del yo, y con ello su solución al problema del libre albedrío, es bastante cuestionable.

Cuando presentamos la crítica que hace Munévar a Llinás sobre el concepto de abstracción, observamos que si bien es cierto que las cualidades secundarias de los sentidos son invenciones de nuestro cerebro o SNC, de ahí no se sigue que los objetos de percepción sean objetos abstractos. Así, retomando el ejemplo anterior, el color es una propiedad que abstrae nuestro cerebro, pero el objeto, las partículas a las cuales le atribuimos el color rojo no son inventadas por nuestro cerebro. Así como el cuaderno y nuestra percepción del cuaderno no son lo mismo, para Munévar, nuestra percepción del yo también debería considerarse diferente del yo (*La evolución*, 260).

Pero, ¿a qué nos referimos con percepción del yo? Y si aceptamos esta diferenciación ¿cuál es entonces el ‘verdadero’ yo?

La percepción de nuestro yo no es otra cosa que la conciencia que tenemos de él. Cuando nos hacemos conscientes de nuestras acciones y deliberaciones estamos percibiendo cómo actúa nuestro yo. “Cuando nos hacemos conscientes de nosotros mismos solamente tenemos una manera simple y conveniente de percibir un aspecto de nuestra muy compleja interacción neuronal” (Munévar, “A Darwinian” 8). La conciencia, lo que en la tradición se ha denominado yo consciente, es solo una percepción del funcionamiento de nuestro cerebro, de nuestro ‘sí mismo’. El cerebro constituye (es) nuestro yo y la conciencia es solo una percepción, un aspecto, de las múltiples relaciones e interacciones neuronales que ocurren en él.

Esta manera de ver el problema permite escapar a una fuerte objeción en contra del libre albedrío basada en varios experimentos que parecen mostrar que el ‘yo consciente’ realmente no causa ni determina ninguna acción. Dentro de estos experimentos se destaca el del neurólogo estadounidense Benjamin Libet. Libet realizó un experimento para tratar de determinar el tiempo requerido para que una decisión consciente cause el movimiento de una mano. A primera vista, nuestro proceso de decisión para, por ejemplo, mover una mano tendría que estar constituido por tres etapas: la decisión consciente del individuo, una señal enviada por el cerebro después de haber sido consciente de dicha decisión y, finalmente el movimiento de la mano. El experimento de Libet dio un resultado distinto y sorprendente: “los humanos se hacen conscientes de la intención de actuar de 350 a 400 ms. luego de que se envía el RP [potencial de preparación], pero 200 ms. antes de que se produzca la acción” (“Do We Have” 47). Esto parece indicar que la conciencia no puede causar mi decisión de mover la mano, pues nos hacemos conscientes de las decisiones después de que han sido tomadas por el cerebro.

Según Munévar, de mantener una identificación entre la conciencia y el yo determinante de nuestras acciones, esta conclusión indicaría que nuestro libre albedrío es tan solo una ilusión, pues nos hacemos conscientes de nuestras decisiones 200 ms. después de que han sido tomadas. En este sentido, una diferenciación entre la conciencia como percepción del yo y el cerebro como yo permitiría afirmar que la ilusión de nuestra elección consciente constituye una ilusión de una percepción, así como la ilusión de que el cuaderno que se encuentra encima de mi escritorio es rojo.

Para Munévar, el cerebro en su complejidad es el que determina nuestras acciones, *el cerebro es el sí mismo*. Nuestra actividad cerebral produce procesos conscientes y no conscientes y ambos son determinantes de nuestras acciones. “La mayor parte de nuestra vida mental no es consciente y, por consiguiente, los ‘archivos’ [la ‘experiencia’, podríamos decir,] del organismo individual los lleva el cerebro cuando ejecuta sus funciones afectando y combinando en su mayoría procesos no conscientes” (Munévar, “A Darwinian” 262-263). Esta visión permite brindar una explicación distinta de los resultados del experimento de Libet: es un proceso inconsciente de nuestro cerebro/yo el que causa la decisión. Pero, si tenemos en cuenta que las relaciones neuronales de nuestro cerebro funcionan de acuerdo con unas leyes fisicoquímicas que consideramos deterministas ¿no estamos cerrando el espacio para el libre albedrío? ¿Las acciones determinadas por nuestro cerebro/yo no están, después de todo, determinadas por esas leyes fisicoquímicas?

Para responder a estos cuestionamientos, Munévar apela a la complejidad del cerebro que, aunque es un órgano como cualquier otro, es el más complejo que tenemos. Como bien lo afirma Churchland, el cerebro es un sistema “en el cual, al menos ocasionalmente, incluso la diferencia más pequeña en su estado actual será rápidamente ampliada en diferencias mucho más grandes en los estados subsecuentes” (*The Engine* 113). Fruto de esa complejidad, el cerebro posee ‘procesos emergentes fuertes’, es decir, procesos en el nivel sistémico que afectan al sistema mismo<sup>4</sup>. “Un estado neurológico es emergente en el sentido en que las conexiones sinápticas que lo constituyen no son suficientes para determinarlo y también emergente en el sentido de que los pesos sinápticos son parcialmente dependientes del estado neurológico mismo” (Munévar, *La evolución* 275). Para comprender esto, recurramos a un ejemplo presentado por John Searle:

Imaginemos una rueda que avanza cuesta abajo. La rueda está íntegramente constituida por moléculas. El comportamiento de las moléculas produce la característica de nivel superior o sistémico de la solidez. Obsérvese que la solidez afecta el comportamiento de las moléculas individuales. La trayectoria de cada molécula se ve afectada por el comportamiento de la rueda sólida

---

<sup>4</sup> En términos neurobiológicos esto se explica porque “los sistemas o redes del cerebro consiguen estados de estabilidad temporales ‘rotando por’ muchos pesos sinápticos tentativos y ajustándolos para lograr una acomodación adecuada a los objetivos del momento” (Munévar, *La evolución* 275).

considerada como un todo. Pero está claro que no hay aquí nada más que moléculas. De modo que, cuando decimos que la solidez influye causalmente en el comportamiento de la rueda y en el de las moléculas individuales que la componen no estamos diciendo que la solidez es algo añadido a las moléculas, sino que es simplemente la condición en que estas se hallan. Pero la característica de la solidez es, no obstante, una característica real que tiene efectos causales reales. (*Libertad* 43-44)

Tenemos algo en el nivel superior, sistémico, que no es ontológicamente reductible a las partes que componen el sistema, pero que no existiría si todas esas partes no se juntaran de una determinada manera. Asimismo, esta propiedad sistémica tiene la capacidad de modificar las partes y al sistema mismo.

La posición de Munévar es, entonces, que el cerebro constituye un sistema emergente fuerte,

un sistema que, aunque usa las leyes de la física y la química, añade sus propios modos de operación, sus propias leyes, para transformar la ambigua información externa en su mundo. Este sistema *sui generis*, el cerebro como 'sí mismo' determina entonces qué acción le parece apropiada. (Munévar, *La evolución* 275)

El cerebro/yo interpreta la información externa como motivo o razón para realizar una u otra acción y la incluye dentro de un conjunto de interpretaciones que previamente posee. Por ejemplo, "una mirada hostil de un extraño me motiva a buscar rápidamente un objeto pesado que pueda agarrar con la mano, pero así sucede sólo porque interpreto su mirada como hostil y porque leo la situación como peligrosa" (Munévar, *La evolución* 249). En este caso, el cerebro/yo interpreta una situación externa, la coteja con información interna que posee, esta se convierte en un motivo y, de acuerdo con *sus* leyes emergentes, decide realizar una determinada acción. Este proceso de deliberación es un proceso inconsciente.

Ilustremos la posición de Munévar considerando el siguiente ejemplo: en un tiempo t1 experimento un sentimiento de sed y en un tiempo t2 me tomo una soda. Lo que afirma este autor es que el estado neurológico que mi cerebro tiene en t1 no es causalmente suficiente para la acción que voy a llevar a cabo en t2. Pero esto no implica que entre t1 y t2 haya

total aleatoriedad, sino que, en ese lapso, el cerebro/yo, por intervención de sus propias leyes, inconscientemente hace la deliberación y toma una decisión. Para decidir mi acción en t2 mi cerebro/yo interpreta los datos del mundo exterior de acuerdo con los datos de 'su mundo interior' y, gobernado por sus propias leyes, toma una decisión de la que ulteriormente me hago consciente. Esas leyes propias del cerebro tienen un estatus diferente de las leyes fisicoquímicas, son *sui generis*.

Pero ¿hasta qué punto soluciona la posición de Munévar el problema del libre albedrío? A nuestro modo de ver, esta posición abre más preguntas de las que soluciona. Empecemos por cuestionar las leyes *sui generis*. Munévar introduce estas leyes para explicar cómo el cerebro/yo determina las acciones de acuerdo con sus propios motivos. Lo único que sabemos de estas leyes es que son propias de los procesos complejos del cerebro y son distintas de las leyes del mundo. A decir verdad, el recurso al cerebro como un sistema emergente fuerte que toma decisiones inconscientemente gobernado por unas leyes propias, con un estatus distinto de las 'leyes naturales' parece completamente *ad hoc*.

Munévar parece estar a favor del compatibilismo, según el cual, libertad y determinismo no son excluyentes (*La evolución* 224). Siendo así, ¿por qué incluir leyes *sui generis* si no se desea contravenir las leyes deterministas fundamentales de la física? Parece que solo tiene sentido afirmar que el cerebro actúa según leyes *sui generis* si se desea escapar al determinismo, pero si lo que se desea es demostrar la compatibilidad de este con la libertad ¿para qué leyes de este tipo? Estas leyes no dejan de ser misteriosas y un recurso explicativo *ad hoc*.

Aun si aceptamos estas leyes propias del cerebro, esto no contribuye a la solución del problema del libre albedrío, pues habría que mostrar cómo funcionan esas leyes, sobre todo, ante los diversos cursos de acción. Podemos aceptar que el cerebro/yo, de acuerdo con leyes *sui generis* sea el determinante de sus acciones, pero es necesario dar una respuesta concreta al problema de las posibilidades alternativas de acción<sup>5</sup>.

Los estudios neurobiológicos del cerebro conducen a Munévar a una conclusión distinta sobre el yo: el yo se puede extender al cerebro en su

---

<sup>5</sup> Con respecto a la responsabilidad moral, la propuesta de Munévar también posee problemas. El principal de ellos radica en del yo al inconsciente: ¿cómo podemos hacernos responsables por haber tomado una decisión de la cual no fuimos conscientes?

complejidad y, con esto, abarcar tanto los aspectos conscientes como lo inconscientes de la actividad cerebral que actúa de acuerdo con unas leyes *sui generis*. Como observamos, esta propuesta tiene sus propios problemas, el mayor de ellos tiene que ver con la naturaleza de las leyes que se postulan para dar respuesta al problema del libre albedrío. En la siguiente sección observaremos otra manera de entender el yo en relación con el cerebro y sus consecuencias para este problema.

### Searle: la naturaleza del yo y el libre albedrío

John Searle presenta un argumento trascendental con el que pretende mostrar la necesidad de postular la existencia de un yo irreductible y de carácter no-humano para explicar las acciones (entre ellas las acciones libres). El carácter irreductible del 'sí mismo' radica en que su existencia no se puede reducir a neuronas o partículas físicas (ontología de tercera persona), sino que tiene una existencia 'subjética' (ontología de primera persona)<sup>6</sup>. El carácter no-humano radica en que su existencia no se limita a un 'lugar' donde se cruzan todas las percepciones y, además, posee un carácter ejecutivo, esto es, tiene injerencia en el mundo físico. Antes de analizar el argumento trascendental con respecto al yo que presenta Searle, es necesario recordar algunas de sus observaciones acerca del problema del libre albedrío.

Como ya dijimos, Searle enuncia de manera clara la tensión que existe entre nuestro sentimiento de libertad y nuestra creencia acerca del funcionamiento determinista del mundo. Pensamos que las explicaciones de los fenómenos naturales han de ser completamente deterministas y, al mismo tiempo, cuando tratamos de explicar nuestro comportamiento, parece que tenemos la experiencia de actuar libre y voluntariamente, de manera que ninguna explicación determinista puede dar cuenta de nuestros actos (Searle, *Libertad* 29-30). Este sentimiento de libertad, según Searle, se basa en la existencia de brechas entre las razones por las cuales un agente toma sus decisiones y la decisión que realmente toma (*Rationality* 13).

Searle reconoce que existen al menos tres brechas: la primera se encuentra en la toma racional de decisiones y radica en que no hay una continuidad

---

<sup>6</sup> "La conciencia tiene una ontología de primera persona; esto es, sólo existe *como experimentada* (*experienced*) por algún humano o animal y, por consiguiente, no puede ser reducida a algo que tenga una ontología de tercera persona, algo que exista independientemente de las experiencias. Es tan simple como eso" (Searle, "Why I am not" 60).

clara entre las razones para tomar una decisión y la decisión. Así, regresando al ejemplo del apartado anterior, uno puede sentir mucha sed y tener la creencia de que la soda quita la sed, pero esto no garantiza (causalmente) que mi decisión sea la de tomarme una soda. La segunda brecha se encuentra entre la decisión y la acción; mi decisión pudo haber sido tomarme una soda, pero esto no implica (no es causa de) que efectivamente mi acción sea tomármela. La tercera brecha se encuentra en las acciones que se extienden en el tiempo, en este caso se puede iniciar la acción, pero esto no implica (no es causa de) que la vaya a terminar (Searle, *Rationality*, 14-15).

Para entender lo anterior consideremos los siguientes enunciados (1) he tomado una soda porque me dio sed, (2) me dio sed porque se me secó la garganta. A primera vista, (1) y (2) tienen la misma estructura lógica (causal), no obstante, el 'porque' en cada uno tiene una connotación distinta. (2) tiene la forma 'A causó B' (la sed causó que mi garganta se reseca). La reseca en la garganta es una causa suficiente para mi sentimiento de sed. En este tipo de explicaciones (causales estándar), el acontecimiento descrito por la oración que precede al 'porque' ocurre una vez dado el acontecimiento descrito a continuación del 'porque' (Searle, *Libertad* 48).

Claramente 'A causó B' no es la estructura lógica de (1), pues como ya dijimos, la sed no es una causa suficiente para tomarme una soda. La explicación de (1) exige postular un yo (un *ego*) que realice la acción. "La forma lógica del enunciado (1) sería: "un yo S ejecutó la acción A y, en la ejecución de A, S actuó por la razón R" (Searle, *Libertad* 50). La explicación de nuestros actos, en este sentido, exige postular la existencia de un 'yo ejecutivo', que determine las acciones<sup>7</sup>. El argumento de Searle puede verse como sigue:

Premisa 1: las explicaciones basadas en razones no suelen indicar condiciones causalmente suficientes.

Premisa 2: dichas explicaciones [las basadas en razones] pueden explicar adecuadamente las acciones.

Premisa 3: las explicaciones causales adecuadas indican condiciones que, en relación con el contexto, son causalmente suficientes.

Conclusión 1: entendidas como explicaciones causales

---

<sup>7</sup> Puede verse este argumento con mayor detalle en Giraldo (*Intencionalidad*).

ordinarias las explicaciones sobre la base de razones son inadecuadas.

Conclusión 2: las explicaciones basadas en razones no son explicaciones causales ordinarias. Aunque tienen un componente causal, su forma no es 'A causó B

Conclusión 3: las explicaciones basadas en razones son adecuadas puesto que explican por qué un yo actuó de una determinada manera. Explican por qué un yo racional actuando en la brecha de la que hablábamos actuó de una manera y no de otra, al especificar la razón a partir de la cual actuó dicho yo. (Searle, *Libertad* 51-54)

El argumento con el cual Searle pretende mostrarnos la necesidad de la existencia de un yo no-humano para la explicación de las acciones, es trascendental. Si aceptamos la existencia de un yo como determinante de las acciones, estamos aceptando eo ipso que existen acciones libres<sup>8</sup>.

Ahora bien, nuestro sentimiento de libertad se basa en la experiencia consciente que tenemos del fenómeno de la brecha, de manera que una de las características fundamentales que, según Searle, posee el yo es que es consciente. El problema del libre albedrío remite así al problema del yo y este, a su vez, al problema de la conciencia. Pero ¿qué es realmente la conciencia para John Searle? En varias partes de su obra él es enfático en afirmar que la conciencia es un fenómeno biológico. Antes de explicar qué significa esto, parece prudente echar un vistazo a qué no es la conciencia para el americano. En *Consciousness and Language* Searle afirma que la conciencia no debe ser confundida ni con el conocimiento ni con la atención (8)<sup>9</sup>.

En primer lugar, los estados conscientes de ansiedad o nerviosismo no tienen ninguna conexión esencial con algo que podamos llamar conocimiento. En segundo lugar, aunque, a menudo, 'ser consciente de' se usa de manera análoga a 'poner atención a', en filosofía de la mente estos dos términos no significan lo mismo. Es necesario distinguir lo que ocurre en el centro de la conciencia, aquello sobre lo que ella

---

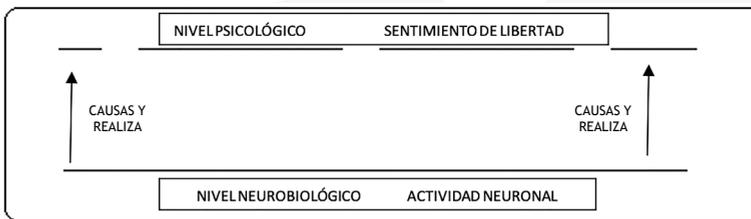
<sup>8</sup> Este argumento también nos permite aceptar que el agente es responsable por las acciones que decidió. "Si asumimos la existencia de una conciencia irreductible actuando sobre la base de razones bajo las restricciones de la racionalidad y la presuposición de que tenemos libertad podemos ahora tener un sentido de la *responsabilidad* y de todas sus nociones operativas. Debido a que aquella misma opera en la brecha sobre la base de razones para tomar decisiones y realizar acciones, éste es el locus de la *responsabilidad*" (Searle, *Rationality* 89).

<sup>9</sup> Así mismo, también afirma que intencionalidad y conciencia no son lo mismo.

está enfocada, de lo que ocurre en su periferia. Así, por ejemplo, en el momento en el que alguien de manera concentrada lee este texto, no está poniendo atención al sentimiento que le produce el reloj apretado en su mano o los calcetines en los pies, esto ocurre porque solo ponemos atención a aquello que está en el centro de la conciencia<sup>10</sup>.

Para explicar el fenómeno de la conciencia recordemos el ejemplo en el que imaginábamos una rueda que avanza cuesta abajo y que, para poder hacerlo, debe estar íntegramente constituida por moléculas. Por medio de este ejemplo Searle ilustra las relaciones existentes entre la conciencia y el cerebro. No hay en el cerebro más que neuronas (así como en la rueda no hay más que moléculas) y, sin embargo, la conciencia no está ubicada en ninguna de ellas ni en ninguna parte del cerebro. La conciencia es una propiedad sistémica del cerebro; la relación dinámica de las partes del cerebro permite que surja la conciencia como una característica del todo. El comportamiento de las neuronas *causa* la propiedad sistémica de la conciencia. Esta última, al igual que la solidez de la rueda, tiene injerencia sobre las partes del todo; en otras palabras, “la conciencia propia del cerebro puede producir efectos en el plano neuronal, aunque no haya en el cerebro nada más que neuronas” (Searle, *Libertad* 44).

Figura. 1



Brechas en los niveles psicológico y neurobiológico

<sup>10</sup> Hay quienes piensan que los fenómenos que hemos descrito son inconscientes. Según Searle esto es un error y la prueba de ello es que podemos pasar rápidamente a prestarles atención, a enfocarnos sobre ellos, hay un cambio de centro y periferia en el campo de la conciencia. (Searle, *Consciousness* 13).

En este gráfico se muestra cómo el nivel neurobiológico causa el nivel psicológico<sup>11</sup>, que a su vez, tiene la propiedad de afectar el primero. También se observa cuál es la naturaleza de nuestro sentimiento de libertad. La pregunta que surge aquí, y que ha planteado Carlos Patarroyo es: “si los estados conscientes son producidos por procesos neurobiológicos, y a nivel psicológico la brecha es real, ¿existe también una brecha a nivel neurobiológico?” (*Freedom* 167). Frente a esta pregunta Searle plantea dos hipótesis, cada una con su correspondiente consecuencia para el problema del libre albedrío.

En la hipótesis 1 podemos pensar que la brecha solo existe en el nivel psicológico; que en el nivel neurobiológico no existe ninguna ruptura, sino que hay una linealidad. El Gráfico 1 nos representa muy bien esta hipótesis. Según esta, el sentimiento de libertad que tenemos solo se representa en el nivel psicológico, en el nivel neurobiológico no existe ninguna libertad, sino que todo se encuentra determinado. Regresemos a nuestro ejemplo: en un tiempo t1 experimento un sentimiento de sed y delibero sobre la posibilidad de tomarme una soda o no hacerlo, finalmente en un tiempo t2 decido tomarme la soda. En el nivel psicológico el sentimiento experimentado en t1 no es una condición suficiente para realizar la acción en t2, es decir, hay un sentimiento de que hubiera podido decidir no tomarme la soda. Sin embargo, según esta hipótesis, ese sentimiento de libertad es solo una ilusión, pues, finalmente, en el nivel neurobiológico no hay ninguna brecha, en este nivel todo estaría determinado. El estado neurobiológico de mi cerebro en t1 es causalmente suficiente del estado neurobiológico de mi cerebro en t2. La brecha psicológica queda explicada porque no somos conscientes de las conexiones que hay entre un estado psicológico y el otro y, por eso, creemos (desde el nivel consciente) que un estado no es causalmente suficiente para el siguiente, cuando en realidad sí lo es. Así, en palabras de Searle, “tendríamos una especie de determinismo neurobiológico que correspondería a un libertarismo psicológico” (*Libertad* 63).

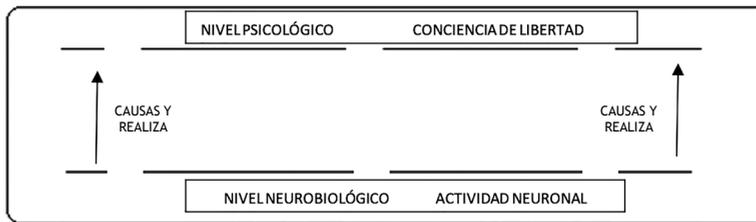
Otra consecuencia de esta hipótesis es la consideración de la conciencia como un epifenómeno que no tiene ninguna injerencia en la decisión. Si

<sup>11</sup> Aquí es necesario realizar una aclaración. Recordemos que Hume ya ha mostrado la independencia lógica que hay entre causa y efecto. Si nos quedamos solo con la palabra ‘causa’ tendríamos que aceptar que el nivel neuronal es *diferente* del psicológico. Por eso, la relación entre estos dos niveles está mediada por la expresión ‘causa y realiza’ para mostrar que no se trata de una causa que no se relaciona más con su efecto, sino que, de alguna manera, es ella misma su efecto. Lo que causa es una configuración de sí misma que produce en sí misma el efecto deseado.

bien es cierto que tenemos una ilusión de que tomamos decisiones con base en razones, nuestras decisiones se encuentran ya determinadas por nuestros estados neurobiológicos, de manera que las razones que posee nuestra conciencia para elegir uno u otro curso de acción no son más que ilusiones, no tienen ninguna eficacia causal real en el mundo. Searle cuestiona los resultados de esta hipótesis, no de manera tan contundente como quisiéramos, abogando que va en contra de todo lo que sabemos de la evolución, pues no parecería muy lógico que nuestro organismo gastara toda esa energía para hacer emerger un aspecto que finalmente no tiene ninguna injerencia en nuestra vida.

Por otro lado, la hipótesis 2 consiste en afirmar que la brecha existe tanto en el nivel psicológico como en el neurobiológico. Para representarla podemos usar el siguiente gráfico:

Figura. 2



Brechas en los niveles psicológico y neurobiológico

Aquí observamos que las brechas del nivel psicológico se corresponden con brechas en el nivel neurobiológico. Recordando el ejemplo de la rueda, la conciencia es causada solo por fenómenos neurobiológicos, pero a la vez tiene poder causal sobre ellos. Según Searle, para que esto sea posible el yo, además de ser consciente, debe tener la capacidad de adoptar y ejecutar decisiones en las brechas, es decir, es el que desde la brecha determina la acción (Searle, *Libertad* 81).

La hipótesis 1 parece estar más acorde con las investigaciones que hasta el momento tenemos del cerebro. Basta con observar semejanzas entre esta hipótesis y la posición de Llinás. Sin embargo, al concluir que la libertad es una ilusión, nos deja el sinsabor de que finalmente no actuamos de manera libre en el mundo, sino que todas nuestras acciones están determinadas desde el nivel neurobiológico. Sin una objeción

contundente a esta hipótesis, movidos más por la esperanza de que nuestro sentimiento de libertad sea algo más que una ilusión, aparece la hipótesis 2 en la que se salva la noción de libre albedrío, pero se tiene que aceptar que en el nivel neurobiológico existen brechas. De ser así, y aquí volvemos a usar las preguntas de Patarroyo, “¿cómo puede un proceso fisicoquímico tener un desarrollo causal en donde la causa antecedente no determine suficientemente el estado siguiente?” (*Freedom* 167).

Según Searle, la única manera de tomar en serio la hipótesis 2 es acudiendo a los descubrimientos de la mecánica cuántica. Si la mecánica cuántica es la única forma de indeterminismo real en la naturaleza y si en verdad hay brechas en la realidad empírica, entonces ella ha de entrar en la explicación de la manera como la conciencia actúa sobre los cuerpos (Searle, *Libertad* 53). Como menciona Patarroyo, este argumento falla porque, en primer lugar, supone que la brecha empírica es real y, en segundo lugar, aún si aceptamos que la mecánica cuántica tiene un lugar en la explicación de la conciencia, no es muy claro cuál es, ni mucho menos en qué sentido sería beneficioso para el problema de la libertad (*Freedom* 169). Recordemos que la libertad se destruye cuando no hay autodeterminación, es decir, cuando la voluntad está determinada por algo externo, pero también cuando no está del todo determinada.

Sin embargo, y pese a esta última dificultad, Searle nos permite llegar un poco más allá que Munévar porque con su explicación del fenómeno de las brechas parece que habría cabida a una explicación de la libertad, por medio del argumento trascendental que da del yo. Al ser así podemos observar ahora que las posiciones de Searle y Munévar tienen menos cosas en común de las que a primera vista parecieran tener. Si bien es cierto que ambos autores están pensando en una solución al problema del libre albedrío que se encuentre de acuerdo con los avances de la ciencia y, en este sentido, apelan a dos niveles en su explicación de la mente (el nivel de lo físico o neurobiológico y el nivel de lo ‘emergente’), los caracterizan de manera bastante distinta. Searle dice que las propiedades de lo emergente (en este caso la conciencia) tienen efectos en lo básico (el comportamiento neuronal), pero no afirma que lo emergente deba tener sus propias leyes que puedan transgredir o violar las leyes de la física. Searle acude a la indeterminación cuántica en el cerebro, en el nivel neurobiológico, para que el funcionamiento del nivel de lo emergente, la conciencia, pueda realmente acceder a cursos alternativos de acción o de decisión y que esto no contravenga ninguna ley del funcionamiento del nivel básico o neuronal. Munévar,

como vimos, propone la existencia de leyes *sui generis* que tienen la capacidad de contravenir las leyes fundamentales de la física, cuya naturaleza es desconocida y misteriosa.

En este recorrido hemos podido ver que el estudio neurobiológico del cerebro arroja tres posiciones diferentes para el problema del yo y el libre albedrío, todas ellas con sus respectivas dificultades. En primer lugar, hay una posición negativa sobre la existencia del yo, de acuerdo con la cual es solo una invención nuestra (Llinás). El problema aquí es que no podríamos dar cuenta del libre albedrío, pues, como hemos insistido, para que este exista es necesario proponer un yo ejecutivo. Ante esta dificultad encontramos dos opciones. La primera, extender la noción del yo, o mejor asumir que el yo es lo mismo que el cerebro (Munevar). El problema con esta alternativa es que, por un lado, extiende al yo más allá de la conciencia y, por el otro, postula leyes *sui generis* para explicar cómo el cerebro toma decisiones, lo cual más que una solución al problema del yo y el libre albedrío aplaza el cuestionamiento sobre cuál es el estatus de estas leyes con respecto al determinismo. La segunda opción es aceptar que el yo no es lo mismo que el cerebro, sino que surge de él. Esta salida permite explicar la libertad a través de la existencia de brechas en la conciencia y, además, permite explicar la toma de decisiones por parte del yo sin apelar a leyes *sui generis*. La principal dificultad de esta alternativa es que no aclara si las brechas psicológicas tienen un correlato con brechas neurobiológicas. Aunque esta salida deja muchos interrogantes abiertos parece tener mayor poder explicativo frente al problema del libre albedrío que la anterior.

## REFERENCIAS

Chisholm, Robert. "Human Freedom and Self" en Gary Watson (ed.) *Free Will*. Oxford: OUP. 24-35, 1982. Print.

Churchland, Paul. *The Engine of Reason, the Seat of the Soul*. Cambridge, Mass: MIT, 1996. Print.

Giraldo, Laura. *Intencionalidad y acciones colectivas. Una lectura desde Searle e Isaacs*. Bogotá: Universidad del Rosario, 2018. Impreso.

Kane, Robert. *A Contemporary Introduction to Free Will*. Oxford: OUP, 2005. Print.

Libet, Benjamin. "Do We Have Free Will?" *Journal of Consciousness Studies* 6 (8-9), 1999, 47-57.

Llinás, Rodolfo. *El cerebro y el mito del yo*. Bogotá: Norma, 2002. Impreso.

Munévar, Gonzalo. *La evolución y la verdad desnuda*. Barranquilla: Universidad del Norte, 2008. Impreso.

"A Darwinian Account of Self and Free Will". Martin Brinkworth and Friedel Weinert (eds.) *Evolution 2.0: Implications of Darwinism in Philosophy and the Social and Natural Sciences*, Dordrecht, Springer, 2012. Print.

Patarroyo, Carlos. "Reseña de 'Freedom and Neurobiology'". *Areté* 19 (1), 2007, 165-172.

Searle, John. *Rationality in action* Cambridge, Mass. MIT, 2001. Print. "Why I Am Not a Property Dualist". *Journal of Consciousness Studies* 9 (12), 2002: 57-64.

*Consciousness and language* Cambridge, Cambridge University Press, 2002. Print.

*Libertad y neurobiología reflexiones sobre el libre albedrío, el lenguaje y el poder político*. Barcelona: Paidós, 2004. Impreso.

*La mente: una breve introducción*. Bogotá: Norma, 2006. Impreso.

Watson, Gary. "Free Will". Jaegwon Kim and Ernest Sosa, (eds.), *A Companion to Metaphysics*, pp. 175-182. Cambridge, Mass.: Blackwell Reference. 1994. Print.

**Como citar:**

Forero-Mora, José Andrés y Giraldo Ceballos, Laura I. "El cerebro, el yo y el libre albedrío: Una discusión entre Llinás, Munévar y Searle". *Discusiones filosóficas*. Jul. 22 (39), 2021: 35-53. <https://doi.org/10.17151/difil.2021.22.39.3>