



## Neurociencias, inteligencia artificial e interioridad humana

Tarcisio Porto Nogueira

Universidad Católica de Córdoba, Argentina  

<https://dx.doi.org/10.5209/resf.99659>

Recibido: 14/12/2024 • Aceptado: 16/07/2025

**Resumen.** Actualmente, hay muchas razones para descartar como un contrasentido la afirmación de una “interioridad humana”. En efecto, el asombroso desarrollo de las neurociencias y la I.A. prometen poder explicar la mente humana, al menos en principio, desde la perspectiva exclusiva del observador (tercera persona). Además, un buen número de filósofos contemporáneos abogan por la abolición de la dicotomía interior-exterior. El presente artículo examinará esos argumentos e intentará rehacer los fundamentos para una nueva afirmación de aquella interioridad perdida. Para tal objetivo, será imprescindible entrar en diálogo con las neurociencias y las ciencias de la computación. Las analogías del “molino” de Leibniz, así como la “habitación china” de John Searle ayudarán en gran medida a ese propósito, lo mismo que la noción de “inconmensurabilidad” de los fenómenos físicos y mentales. Finalmente, presentaremos un modelo de articulación entre realidades de índole tan distintas.

**Palabras clave:** dicotomía interior-exterior; neurociencias; inteligencia artificial; reduccionismos; “argumento de la habitación china”; inconmensurabilidad explicativa.

### EN Neurosciences, artificial intelligence and human interiority

**Abstract.** Currently, there are many reasons to dismiss the claim of a ‘human interiority’ as a nonsense. Indeed, the astonishing development of neuroscience and AI promise to be able to explain the human mind, at least in principle, from the exclusive perspective of the observer (objective, third person perspective). Moreover, a good number of contemporary philosophers advocate for the abolition of the interior-exterior dichotomy. This article will examine these arguments and attempt to rebuild the foundations of that lost interiority. To this end, it will be essential to engage in dialogue with neuroscience and computer science. The analogies of Leibniz’s “mill” and John Searle’s “Chinese room” will greatly aid this purpose, as will the notion of the “incommensurability” of physical and mental phenomena. Finally, we will present a model of articulation between realities of such different kinds.

**Keywords:** Interior-exterior dichotomy, neurosciences, artificial intelligence, reductionism, Chinese room argument, explanatory incommensurability

**Sumario:** 1. Introducción; 2. Circunscripción del término; 3. Disolución de la interioridad; 4. La inconmensurabilidad entre los fenómenos físicos y mentales; 5. “Interior” como negación del espacio; 6. El fenómeno de la significación; 7. Conclusión; 8. Referencias bibliográficas.

**Cómo citar:** Porto Nogueira, T. “Neurociencias, inteligencia artificial e interioridad humana”, *Revista de Filosofía*, 51 (2), 259-271.

## 1. Introducción

El avance y expansión verdaderamente extraordinarios de las neurociencias, por un lado, y de la comúnmente llamada “inteligencia artificial”, por otro, constituyen en la actualidad un hecho innegable y de consecuencias relativamente impredecibles para un futuro muy cercano. Un primer resultado al que arriban ya muchos autores es la adquisición de una comprensión nunca jamás lograda hasta ahora acerca de la mente humana y de nuestra especie en general. Se resaltan, sobre todo, el conocimiento de los procesos neurofisiológicos que se desencadenan en el cerebro y su paralelismo con relatos y experiencias conscientes, y el mismo paralelismo cada vez más asombroso entre eventos en otro tiempo tenidos como exclusivos de nuestra subjetividad y su emulación en el terreno de la informática. Todo ello anima a afirmar una creciente capacidad para describir el fenómeno humano con un lenguaje cada vez más exclusivo de una perspectiva de tercera persona, perspectiva que promete fagocitarse progresivamente a lo que antes era patrimonio específico de una perspectiva de primera persona. Es así como se nos depara hoy un movimiento siempre creciente que parece expulsar definitivamente del terreno de la Filosofía – entendida como análisis riguroso y consistente acerca del ser humano y todos los interrogantes que suscita – lo que en un tiempo ya lejano se denominaba “interioridad”.

En este breve trabajo, presentaremos sólo algunas de las múltiples formas de eliminación o negación de esa “interioridad” como para poner en evidencia dicha supresión progresiva, pero sin detenernos en los pormenores de los pasos demostrativos que cada autor recorre en sus obras, lo que excedería en mucho nuestros límites. A continuación, intentaremos recuperar esa supuesta dimensión del ser humano, eligiendo asimismo otros filósofos – con el mismo procedimiento – cuyas reflexiones abonan la posición contraria. También por motivos de brevedad revisaremos muy rápidamente el reduccionismo de tipo neurocientífico para luego poder detenernos algo más en uno de sus “parientes” más cercanos, basado en la *teoría computacional de la mente*.

## 2. Circunscripción del término

Antes de poder dar un paso siquiera, se hace imprescindible ofrecer de entrada una noción, aunque provisoria y perfectible, de lo que entendemos por “interioridad”. Con esa expresión queremos circunscribir un campo semántico desde el que puedo decir que tomo conciencia de mi cuerpo, del mundo y de mí mismo, campo de sentido supuestamente irreductible a los eventos neurofisiológicos a través de los cuales dichas actividades “interiores”, acciones o procesos pueden constatarse, en alguna medida, desde una perspectiva de tercera persona. Aclaremos también que dicho campo semántico no se reduce a una realidad puramente lingüística, sino que hace referencia a entidades o eventos extralingüísticos como, por ejemplo, yo mismo, mis pensamientos, emociones, conciencia de cualidades sensibles, etc. Así lo expresaba Agustín de Hipona: “No quieras irte afuera. Entra dentro de ti mismo; porque en el hombre interior habita la verdad”<sup>1</sup>.

Una buena ejemplificación de estas vivencias internas la encontramos en una descripción que tomaremos de Edith Stein. En ella advertimos un relato en primera persona, difícilmente reproducible – y menos aún entendible – en tercera persona:

Yo reflexiono sobre un problema difícil y trato en vano de encontrar una solución. En fin, lo abandono, porque “hoy me encuentro muy atontada” (*dumm*). No puedo percibir esa tontería (*Dummheit*) mía a través de mis sentidos externos (hacemos abstracción de los aspectos perceptibles exteriormente que puedan verse en el cuerpo); no puedo tampoco estar “inmediatamente consciente” [de mi tontería], pues sólo la reflexión me la manifiesta. Pero la experiencia, ella se me revela de la misma manera como experimento un cuchillo embotado que rehúsa cortar el pan. La forma original de tal experiencia es el punto de partida de juicios y conclusiones ulteriores y, conservada en la memoria, nos conduce a reunir un conjunto sólido de experiencias que nos permiten conocernos a nosotros mismos. Esta forma la llamamos con Husserl “percepción interna” (*innere Wahrnehmung*)<sup>2</sup>.

En este sentido, el término “vivencias<sup>3</sup>” nos da a entender con más claridad lo que, sobre todo desde la Fenomenología, entendemos por “interioridad”, aunque ésta no se agote en vivencias efectivas, sino que también puede entenderse estados no conscientes, siempre y cuando puedan, en principio, ser experimentados por el sujeto de dichos estados. Finalmente, incluimos dentro de ese campo semántico

<sup>1</sup> Sant’Agostino (1930), p. 264. Cf. Uña Juárez (2001), pp. 31-52.

<sup>2</sup> Stein (1996), pp. 381-382. Hemos hecho algunas modificaciones según el original Steins Werke (1950), p. 337. El contenido entre corchetes será siempre nuestro.

<sup>3</sup> En el § 46 de *Ideas I*, el maestro de Stein habla de la “indubitabilidad de la percepción inmanente y la dubitabilidad de la percepción trascendente” para subrayar el carácter *evidente* de nuestras vivencias: “No hay ningún contrasentido en la posibilidad de que no exista ninguna de las conciencias ajenas que afirmo en la experiencia intrafactiva [percepción interna]. Pero mi intrafacción, mi conciencia en general, está dada originaria y absolutamente, no sólo en esencia, sino en existencia. Sólo para el yo y la corriente de las vivencias en relación consigo mismas se da esta señalada situación; sólo aquí hay algo como una percepción inmanente, y no puede menos de haberla” (Husserl (1993), p. 105).

aquellos *qualia*<sup>4</sup>, es decir, las cualidades conscientes que suelen “acompañar” o seguir a las *sensaciones*, entendidas estas últimas en su sentido enteramente neurofisiológico.

### 3. Disolución de la interioridad

Ahora bien, el hecho de que dichas vivencias constituyan un ámbito absolutamente privado del sujeto que las experimenta – hecho el más obvio y evidente para la Fenomenología – ha sido cuestionado e incluso rechazado como ilusorio por gran parte de la filosofía contemporánea, sobre todo aquellas corrientes que celebran la superación del así llamado “paradigma del sujeto” o de la conciencia. Y el rechazo más o menos vehemente de cualquier “interioridad” o “espiritualidad” constitutiva del ser humano no es nada original de nuestro tiempo. Al menos en ambientes académicos, dicha impugnación ya se anunciaba explícitamente a mediados del siglo pasado:

El mundo no está ahí como un objeto de cuya constitución poseo la ley, sino que es el medio natural y el campo de todos mis pensamientos y de todas mis percepciones explícitas. La verdad no “habita” solamente el hombre interior<sup>5</sup> o, más bien, *no hay hombre interior*; el hombre está en el mundo, es en el mundo que se conoce a sí mismo. Cuando vuelvo a mí dejando el dogmatismo del sentido común o el de la ciencia, encuentro, ya no un hogar de verdad intrínseca, sino un sujeto entregado (*voué*) al mundo<sup>6</sup>.

Más allá de si podemos o no creerle bajo palabra a Merleau-Ponty, lo cierto es que en este pasaje aboga ya por una abolición de la interioridad. Si bien, en el entramado de su discurso, quedará un “resto” de la misma, esa subjetividad restante pasará a ser parte inseparable del mundo – y viceversa – , en una “inextricable confusión”<sup>7</sup>. Sea como fuere, lo que nos interesa constatar aquí es el testimonio de un creciente descrédito e incluso la supresión de ese *homo interior* – para seguir con la expresión agustiniana, y designar mediante ella ese ámbito privado, sólo experimentable y descriptible en primera persona – .

Once años más tarde, y siguiendo coherentemente el impulso desintegrador de “Las palabras y las cosas” (1966), Michel Foucault constata la omnipresencia del lenguaje con la desaparición del sujeto y de su conciencia, juntamente a una nueva impugnación de la interioridad:

La transición hacia un lenguaje en que el sujeto está excluido, la puesta al día de una incompatibilidad, tal vez sin recursos, entre la aparición del lenguaje en su ser y la conciencia de sí en su identidad, es hoy en día una experiencia que se anuncia en diferentes puntos de la cultura (...). Este pensamiento que se mantiene fuera de toda subjetividad para hacer surgir como del exterior sus límites, enunciar su fin, hacer brillar su dispersión y no obtener más que su irrefutable ausencia, y que al mismo tiempo se mantiene en el umbral de toda positividad, no tanto para extraer su fundamento o su justificación, cuanto para encontrar el espacio en que se despliega, el vacío que le sirve de lugar, la distancia en que se constituye y en la que se esfuman, desde el momento en que es objeto de la mirada, sus certidumbres inmediatas, –este pensamiento, con relación a la interioridad de nuestra reflexión filosófica y con relación a la positividad de nuestro saber, constituye lo que podríamos llamar en una palabra “el pensamiento del afuera”<sup>8</sup>.

De una forma, diríamos, mucho más ordenada y sistemática, Richard Rorty en los Estados Unidos dará el “golpe de gracia” a la quimérica “interioridad humana” mostrándonos tanto su origen como su fin en la historia del pensamiento. Rorty empieza por cuestionar certeramente la noción de “idea” en Descartes y Locke, para terminar revocando en duda el “teatro cartesiano”:

La novedad estuvo en la idea de un solo *espacio interior* en el que eran objeto de cuasi observación las sensaciones corporales y perceptivas (...), las verdades matemáticas, las reglas morales, la idea de Dios, los talentos depresivos, y todo el resto de lo que llamamos “mental”. Este escenario interior, con su *observador interno*, había aparecido insinuado en varios momentos del pensamiento antiguo y medieval pero nunca se había tomado en serio el tiempo suficiente como para servir de base a una problemática<sup>9</sup>.

Una cosa es cierta: Descartes en ningún momento cuestiona la interioridad de la *res cogitans*. Ni siquiera en la *Meditación primera*, donde la duda metódica adquiere su forma hiperbólica, ni siquiera allí, decimos, las *cogitationes* son puestas en tela de juicio; sí, los correlatos “externos” de las mismas. Tanto es así que los correlatos de las ideas de sensación Descartes los denomina “cosas materiales” o “cosas exteriores”, precisamente en contraste con la *interioridad* de los pensamientos. Por tanto, en las antípodas de Rorty, Foucault y Merleau-Ponty, Descartes, al comienzo de sus *Meditaciones* sólo admite esa interioridad, suponiéndola todo el tiempo, como algo que no necesita demostración<sup>10</sup>.

<sup>4</sup> Cf. Goldstein (2020), pp. 32-33.

<sup>5</sup> “*In te redi; in Interiore homine habitat veritas*” (Saint-Augustin). Nota del original (*sic*).

<sup>6</sup> Merleau-Ponty (1945), p. V (énfasis nuestro).

<sup>7</sup> Merleau-Ponty (1945), p. 518; cf. Bech (2005), pp. 15, 250ss; Porto Nogueira (2016), pp. 11-24.

<sup>8</sup> Foucault (2004), pp. 09-10. El original francés es de 1966, el mismo año de *Las palabras y las cosas*. Cf. también Castro (2011), pp. 172 y 209.

<sup>9</sup> Rorty (1989), pp. 54-55.

<sup>10</sup> San Agustín “recoge el guante” de los escépticos también respecto del “yo” y sus pensamientos, y les propone en varias de sus obras una solución. En esto se muestra todavía más insistente que Descartes: cf. Uña Juárez (2001), pp. 44-45.

Mas volviendo a Rorty – que al negar esa supuesta dimensión humana devuelve el *onus probandi* a los que la postulan – concluyamos por ahora con él:

(...) ya no nos sentiremos tentados por más tiempo por la noción de que el conocimiento es posible gracias a una Esencia de Vidrio especial que permite a los seres humanos reflejar la naturaleza. De esta manera no sentiremos la tentación de pensar que la posesión de una vida interior, una corriente de la conciencia, tiene relación con la razón<sup>11</sup>.

Notemos que el filósofo estadounidense no se refugiará en alguna especie de irracionalismo, y por eso tiene el cuidado de distinguir la ya malhadada “vida interior” de la razón. Una razón, por cierto, que abandona cualquier indicio de fundamento objetivo e incommovible, precisamente porque confía en un diálogo completamente abierto a todas las posibilidades del consenso. En efecto, para Rorty la única función constructiva que le queda aún a la Filosofía – y no privativa de ella – es la *conversación (conversation)*<sup>12</sup>, actitud dialogante siempre constructiva y *edificante*, que depone la pretensión de universalidad, objetividad acultural y sistematicidad; en una palabra, que desiste para siempre de intentar reanudar una vez más una teoría del conocimiento (*Epistemology*)<sup>13</sup>.

Pasando ahora a otro terreno de discusión, este intento de cancelación de la interioridad se verá en gran parte ratificada por la palabra cada vez más autorizada en nuestros días de las neurociencias, y eso no viene de ahora. Desde décadas atrás, las neurociencias – o, más precisamente, muchos filósofos de la mente apoyado en ellas – se proponen eliminar cualquier “interioridad” y de tal manera que no se hable más de “mente”, sino de cerebro. Un ejemplo claro, sistemático y con una solidez argumentativa impresionante es el del también estadounidense Daniel Clement Dennett (1942-2024).

En su libro *La conciencia explicada* (1991), y luego de haber provisionalmente eliminado la perspectiva de primera persona para permanecer en un “terreno neutral”<sup>14</sup>, Dennett entiende la conciencia como un mero “producto biológico”. Así lo expresa, con su estilo claro y siempre entretenido:

(...) las construcciones más extrañas y maravillosas de todo el mundo animal son las increíbles y complejas construcciones que levanta un primate, el *Homo sapiens*. Todo individuo normal de esta especie construye un yo. A partir de su cerebro teje una tela de palabras y de actos y, como las demás criaturas, no tiene por qué saber qué está haciendo; sólo lo hace. Esa tela lo protege, como la concha del caracol, le proporciona el sustento, como la tela de la araña, y favorece sus perspectivas para el sexo, como la glorieta del pájaro jardinero. (...) Ese “tejido de discursos” (...) es tanto un producto biológico como lo son las demás construcciones que encontramos en el mundo animal<sup>15</sup>.

Estrictamente hablando, el yo no es real, sino una ficción extremadamente sofisticada y útil. Por ser tan útil, ha sido cuidadosamente desarrollada en la historia de la evolución para nuestra especie como un “centro de gravedad narrativa”<sup>16</sup> que simplifica enormemente el lenguaje cuando se trata de hacer atribuciones, describir las propiedades de alguien o de mí mismo. La vívida y única impresión que tenemos de ser un “yo” no es más que una ilusión, lo que desde Dennett se ha dado por llamar “la ilusión del usuario”:

Cuando interactúo con el ordenador, tengo un acceso limitado a los acontecimientos que se producen en su interior. Gracias a los esquemas de presentación ideados por los programadores, me veo invitado a participar de una elaborada metáfora audiovisual, una obra interactiva representada en el escenario del teclado, el ratón y la pantalla. Yo, el usuario, estoy sujeto a toda una serie de ilusiones benignas: parezco ser capaz de mover el cursor (un siervo visible y poderoso) hacia el lugar dentro del ordenador donde guardo mi archivo y, cuando veo que el cursor ha llegado hasta “allí”, presionando una tecla recupero el archivo que, como un largo pergamino, se desenrosca frente a una ventana (la pantalla) bajo mi control. Tecleando diferentes mandos y presionando diversos botones, puedo hacer que ocurra todo tipo de cosas en el interior de mi ordenador sin que yo tenga que conocer los detalles; mantengo el control con sólo confiar en mi comprensión de las detalladas metáforas audiovisuales provistas por la ilusión del usuario<sup>17</sup>.

Justamente por ello, nos dice Dennett, se hace tan simple transponer esa metáfora al esquema mente-cerebro: nos dejamos fácilmente ilusionar de que tenemos una mente que nos hace estar en el control aunque ignoremos por completo lo que ocurre al interior de la caja craneana. Sólo nos enteramos de lo que se nos representa en el “teatro cartesiano” de la fenomenología<sup>18</sup>. Entonces, ¿qué? ¿No existimos los seres humanos? Y Ud., lector, ¿no existe? Claro que sí, nos lo confirma Dennett, y hay más:

<sup>11</sup> Cf. Rorty, *op. cit.*, pp. 43-44.

<sup>12</sup> Cf. *Ibid.*, pp. 337ss.

<sup>13</sup> Cf. *Ibid.*, pp. 332ss.

<sup>14</sup> Cf. Dennett (1995), pp. 84-85.

<sup>15</sup> Dennett, *op.cit.*, pp. 426-427.

<sup>16</sup> Cf. *Ibid.*, pp. 439.

<sup>17</sup> Cf. Dennett (1995), pp. 324-325; cf. igualmente: Frankish (2017), p. 20.

<sup>18</sup> Dennett. *op.cit.*, p. 325.

Si usted piensa en usted como un centro de gravedad narrativa, por otra parte, su existencia depende de la persistencia de esa narración (...), que *en teoría* podría sobrevivir a una serie indefinida de cambios de *medio*, podría teletransportarse fácilmente (en principio) como el noticiero de la noche y almacenarse por tiempo indefinido en forma de mera información. Si lo que usted es, es esa organización de la información que ha estructurado el sistema de control de su cuerpo (o, por plantearlo de manera más provocativa y, a la vez, más usual, si lo que usted es, es el programa que corre en el ordenador de su cerebro), entonces, en principio, usted podría sobrevivir a la muerte de su cuerpo tan intacto como un programa que puede sobrevivir a la destrucción del ordenador en el que fue creado por primera vez<sup>19</sup>.

Estamos comprobando, pues, que “La conciencia explicada” de Daniel Dennett nos hizo transitar sin apenas notarlo desde el reduccionismo neurocientífico a un reduccionismo más explícitamente informático. Y como si faltaran más voces a este ingente clamor, los nuevos y asombrosos éxitos de la llamada “inteligencia artificial” (IA) nos tienen a todos maravillados, también a los que nos dedicamos a la Filosofía. La confianza en nuevos progresos en el campo de la informática es tal que nos hace creer que aquellos estados fenoménicos o intencionales podrían ser artificialmente replicados mediante procesamiento de información, ahora potenciados al infinito con el nuevo LLM (*Large Language Model*). Esto sería todavía más factible cuando se disponga de la tecnología suficiente como para conectar la IA al mundo que la rodea mediante sofisticados sensores y “actuadores” (*actuators*). Se habla entonces de una “cognición corporeizada” (*Embodied Cognition*)<sup>20</sup> conseguida en base a esa interacción con el medio, tal y como lo hacemos los humanos.

Aún así, pese a todos los asombrosos avances, se es consciente de la ingente tarea teórica y tecnológica que supone la equiparación de la máquina con el cerebro. “La cuestión clave para el futuro de la IA”, se reconoce, “es [saber] cómo los humanos pueden aprender tanto sólo de la mera observación”<sup>21</sup>...

Pero volviendo al terreno filosófico, sobre todo ante las impugnaciones de Foucault y, más claramente aún, de Rorty, no se vislumbran fácilmente rutas de salida. En efecto, cualquier punto de partida que se tome parece ya suponer lo que se quiere probar.

Más claramente indefensa se nos presenta la Fenomenología ante las mencionadas impugnaciones. Si tomamos los comienzos propiamente husserlianos, la reducción fenomenológica, fiel a la impronta cartesiana, revoca todo en duda – o practica la *epoché* –, quedando como residuo del procedimiento únicamente la conciencia con sus vivencias, y una “conciencia pura”<sup>22</sup>. A todas luces, nuestro ambicionado botón (la interioridad humana) ya estaba supuesto desde el principio, al menos como condición de posibilidad del mundo y del sentido mismo de cualquier “exterioridad”.

Es menester, pues, volver al mismo “campo de batalla” para desde allí poder positivamente examinar y juzgar acerca de tesis tan radicales. Puestos a la obra, sorprendentemente encontramos, incluso en la actualidad, voces discordantes y que pueden ser muy efectivas a nuestro propósito. Una de las más actuales y pertinentes es la de la filósofa alemana Brigitte Falkenburg (1953-), que introduce en el contexto de esta discusión la noción de *inconmensurabilidad*<sup>23</sup>.

#### 4. La inconmensurabilidad entre los fenómenos físicos y mentales

La filósofa y física alemana, en su obra *El mito del determinismo* (2014), abarca significativamente la historia de los innumerables intentos para subsumir los eventos conscientes desde diversas teorías y modelos, pero señalando siempre la insuficiencia e incluso la inconsistencia de algunos de ellos. En el penúltimo capítulo de su obra, ya está en condiciones de presentarnos su tesis: “Los fenómenos mentales son inconmensurables con los fenómenos físicos. No pueden ser aislados, medidos y analizados causalmente mediante los métodos experimentales aplicables a los fenómenos físicos”<sup>24</sup>.

No podemos exponer ni sintetizar en este lugar toda esta valiosa y detallada obra – ni siquiera el mencionado capítulo – por lo cual nos trasladamos a la justificación más inmediata para su postura:

Las neurociencias cognitivas no han logrado actualmente mucho mejor que hace cien años suavizar la diferencia cualitativa entre los fenómenos mentales y físicos. Ni siquiera pueden hacerlo metodológicamente. No logran demostrar más que ciertas correlaciones<sup>25</sup> entre fenómenos físicos y psicológicos; y todos los métodos, desde la psicofísica hasta los relatos sobre daños cerebrales y los procedimientos de imagen en el sentido más amplio, no proporcionan más que una conexión causal débil en las relaciones entre el cerebro y

<sup>19</sup> Dennett. *op.cit.*, pp. 440-441.

<sup>20</sup> Cf. Krauss (2024), p. 222.

<sup>21</sup> Bengio & Hinton (2021), p. 62. Este hecho sorprendente hizo a Noam Chomsky admitir una estructura cognitiva *a priori* y oponerse al conductismo en el célebre debate con Michel Foucault: cf. Chomsky & Foucault (2006), pp. 02-05.

<sup>22</sup> Husserl (1993), pp. 74-75.

<sup>23</sup> Llegamos al conocimiento de esta filósofa por Gabriel (2019), p. 121. El autor vuelve a insistir sobre el concepto de inconmensurabilidad en una obra más reciente: cf. Gabriel (2024), p.103.

<sup>24</sup> Falkenburg (2012), p. 355.

<sup>25</sup> La autora no cita aquí a David Chalmers, pero esta afirmación ha sido defendida por él en diversas ocasiones: cf. Chalmers (2010).

la mente. Todos los métodos experimentales que quieren lograr más, fracasan en el *abismo* que todavía existe entre nuestra experiencia subjetiva y las cosas y eventos objetivos del mundo exterior (*Außenwelt*)<sup>26</sup>.

Obviemos por ahora la expresión “mundo exterior” para volver a retomarla más adelante, y sigamos presentando los argumentos más significativos con los que la autora justifica su tesis. Habiendo señalado los obstáculos teóricos y metodológicos que dificultan y/o impiden muchas veces el estudio de los fenómenos mentales y sus aproximados correlatos neuronales, dice:

El quid de todas estas dificultades es que los fenómenos mentales no pueden aislarse tan bien como los físicos. Ni siquiera de manera aproximada puede medirse qué sensaciones, ideas y otros contenidos de conciencia entran en un experimento de estímulo-respuesta de un sujeto de prueba, y la información que los mismos sujetos ofrecen al respecto tampoco es muy confiable<sup>27</sup>.

Siendo estas dificultades de índole metodológica, nos preguntamos si no podría haber una razón más fundamental que diera cuenta de las mismas. También en Falkenburg encontramos la respuesta, sobre todo cuando aborda el tema del reduccionismo informático:

La pregunta crucial ahora es ésta: ¿puede superarse o sortearse de alguna manera esta inconmensurabilidad de los fenómenos mentales? El concepto de información ciertamente no lo logra. Sí, intenta hacerlo, pero pasando por alto displicentemente el hecho de que la información que emite una computadora tiene poco en común con nuestros contenidos conscientes; y lo poco que tienen en común se debe más a la conciencia de los programadores y usuarios de computadoras que a la capacidad de cálculo basada en hardware. Sin embargo, no veo por ningún lado otros conceptos científicos que establezcan un *punte semántico* sólido entre los fenómenos físicos y mentales<sup>28</sup>.

Llegados a este punto, el carácter de inconmensurabilidad de los fenómenos mentales nos explica el reiterado fracaso de los intentos de una postura reduccionista. Esto, a su vez, nos ha dejado más cerca de la demostración que se nos pedía: la de la afirmación de una *interioridad* humana, por la que se designaría una experiencia individual a la que no puede tener acceso, por principio, nadie más sino yo mismo, y nunca un observador “externo” (con lo cual la expresión “mundo externo” utilizada por Falkenburg empezaría a cobrar sentido como correlativa a *interioridad* o “mundo interno”). De ahí su incompatibilidad con los métodos científicos, esencialmente pensados desde una perspectiva de tercera persona.

Ahora empezamos a reconocer más bien que aquellas *vivencias* caracterizadas y ejemplificadas al comienzo de este trabajo (recordemos la descripción de Edith Stein) constituyen precisa y positivamente una esfera que, pese a todos los encomiables esfuerzos de la ciencia, no puede ser asumida mediante una cadena o una red o un sistema de causas materiales.

Sin embargo, nuestro trabajo no está del todo hecho. En efecto, no nos basta con decir que la ciencia – hasta el momento – no ha podido subsumir en sus explicaciones los fenómenos mentales. En primer lugar, porque esto caracteriza únicamente una imposibilidad temporal y transitoria. En segundo lugar, dado el talante negativo de semejante tesis, quedaría tipificada estrictamente como una falacia de “apelación a la ignorancia” (*ad ignorantiam*), según la cual estaríamos dando como verdadera una proposición por el simple hecho de que no ha sido aún probado que es falsa.

## 5. “Interior” como negación del espacio

Por esa razón, hemos de volver a Falkenburg. De hecho, su minucioso examen apunta también hacia una “falacia mereológica<sup>29</sup>”, la de confundir el todo y sus partes, en la que incurrirían tanto los partidarios de la reducción informática como los de la reducción neurobiológica. Sin embargo, creemos poder seguir avanzando algo más. En efecto, no se trata de señalar y evitar una inadecuada atribución de las propiedades del todo a una de sus partes (como la atribución de pensar al cerebro), sino en negar en absoluto que el cerebro sea una *parte* del sujeto que piensa. Es decir: la conciencia – o el sujeto consciente – no puede ser reducida ni a las partes de un sistema, ni al sistema como un todo, emergente o no, de esas partes. Sencillamente, no tiene sentido forzar una relación espacial – como la que relaciona el todo con sus partes – entre la conciencia y el cerebro, y mucho menos con alguna de sus actividades neurofisiológicas. En una palabra: la *conciencia no es divisible* y, en este sentido muy preciso, Leibniz tenía razón: constituye una

<sup>26</sup> Falkenburg (2012), p. 355. El énfasis será siempre nuestro.

<sup>27</sup> *Ibid.*, p. 354. (Aprovecho para agradecer aquí al Dr. José Ubaldo Fernández que, entre otros muchos aportes, me tiene siempre actualizado con los últimos hallazgos neurocientíficos).

<sup>28</sup> *Ibid.*, p. 365. Si no hay, o todavía no hay, dicha conexión de sentido entre eventos tan diversos, ¿no estaríamos justificados a hablar de emergencia de los fenómenos mentales, como último y acabado estadio evolutivo de nuestra especie? La autora examina esa hipótesis y la descarta muy justificadamente (cf. *Ibid.*, p. 358; cf. Bennett & Hacker (2003), p. 65).

<sup>29</sup> Clásicamente denominada “falacia de división”: cf. Carney & Richard (1964), p. 50; Hamblin (2016), pp. 28ss. La autora la recoge de Bennett & Hacker, *op. cit.*, pp. 68ss, aunque la detecta más allá de un mero error lógico-lingüístico y lo señala concretamente en argumentos que se construyen “a partir de hallazgos neuropatológicos” (cf. Falkenburg (2012), p. 356).

*unidad originaria* (“mónada”). Esto nos obliga a visitar – y será con provecho – uno de los párrafos más famosos de su *Monadología*:

Hay que reconocer, por otra parte, que la percepción y lo que de ella depende resultan inexplicables por razones mecánicas, esto es, por medio de las figuras y de los movimientos. Porque, imaginémosnos que haya una máquina cuya estructura la haga pensar, sentir y tener percepción; se la podrá concebir agrandada, conservando las mismas proporciones, de tal manera que podamos entrar en ella como en un molino. Esto supuesto, una vez dentro, no hallaremos sino piezas que se impelen unas a otras, pero nunca nada con qué explicar una percepción. Así, pues, esto hay que buscarlo en la sustancia *simple*, no en lo compuesto o en la máquina<sup>30</sup>.

No vamos a defender aquí el ingenioso sistema monadológico leibniziano, pero es suficiente entender el término “percepción” (que para Leibniz tiene un sentido técnico muy preciso) como un concepto que abarca todos aquellos estados cualitativos que reconozco experimentar exclusivamente en primera persona (como aquellos descritos por Edith Stein, también en primera persona). Bajo este concepto, la analogía del molino (o de cualquier máquina, por sofisticada que sea) se adecua perfectamente para ilustrar esa imposibilidad de principio que progresivamente veníamos constatando: la de una reducción de la conciencia o de “lo mental” a cualquier realidad material, vale decir, cuantificable.

Que el lenguaje del s. XVIII no nos depare ya un impedimento y nos lleve a descartar rápidamente el argumento de Leibniz como falto de valor. Porque aún en el caso de que leamos el presente párrafo de la *Monadología* bien provistos de nuestra mecánica cuántica, ésta no nos permitiría de todos modos hablar en términos *no espaciales*, y es esto precisamente lo que expresa Leibniz con el calificativo de “simple” (*simplex*): la ausencia de partes. Centenares de años después, la expresión medieval *partes extra partes*<sup>31</sup> sigue teniendo sentido, no importando ya que esas “partes” sean tan diminutas como las del “modelo estándar de partículas”. Por tanto, para poder decir que la conciencia es reductible a un evento material (físico-químico), es perentorio que tenga sentido decir asimismo que la conciencia ocupa un lugar (en el interior de la caja craneana, por ejemplo). Aplicar categorías espaciales a actos representativos o significativos carece a todas luces de sentido, y esto por una especial “índole de la comprobación requerida por su sentido<sup>32</sup>”.

## 6. El fenómeno de la *significación*

En ese estado de cosas, arribamos a un tópico que pide un mayor detenimiento. Acabamos de ver que el fenómeno de la significación o la dimensión *semántica* del lenguaje no se deja encerrar ni reducir a un evento cuantificable o, lo que termina siendo lo mismo, computable. Esto lo han constatado muchos filósofos, además de Leibniz, Husserl y otros aquí mencionados, claro está, y la historia se remonta hasta Platón<sup>33</sup>, por lo menos. Mas por razones de brevedad, pasamos directamente al s. XX, y elegimos a dos filósofos que han explicitado claramente este tema.

Nuestra primera elección recae sobre Georges Canguilhem (1904-1995). Filósofo y médico al mismo tiempo, nos pone de manifiesto que la filosofía no ha de andar reñida con las ciencias, pese a sus respectivas especificidades. El médico filósofo expresaba la irreductibilidad de nuestra capacidad de significación (o semántica) de una forma muy particular, distinguiendo lo que es una “relación entre...” y una “relación con...”:

En síntesis, el lenguaje humano es esencialmente una función semántica, de la que las explicaciones de tipo fisicalista no han logrado nunca dar cuenta. Hablar es significar, dar a entender, porque pensar es vivir en el sentido. El sentido no es *relación entre...*; es *relación con...* (*relation à...*). Esto se debe a que el sentido escapa a toda reducción que intenta localizarlo en una configuración orgánica o mecánica. Las máquinas llamadas “inteligentes” son máquinas para producir *relaciones entre* los datos que se les proporciona, pero las máquinas no están en *relación con* aquello que el usuario se propone [hacer] a partir de las relaciones que ellas producen para él. Puesto que el sentido es *relación con*, el ser humano puede jugar con el sentido, deformarlo, fingir, mentir, tender trampas<sup>34</sup>.

Es sumamente interesante y pertinente la mención de “tender trampas” (*tendre des pièges*). En efecto: nosotros podemos mentir, engañar, embaucar con toda clase de falacias – reservándonos *ad intra* la verdad – precisamente porque, además de admitir esa dimensión “interna”, constatamos asimismo que la

<sup>30</sup> Leibniz (1981), p. 87 (publicada por primera vez el año 1714). Este dato también se lo debemos a Gabriel (2019), p. 121.

<sup>31</sup> Según la acepción clásica del término, en cuanto nota característica de la “materia”, y así desde el medioevo, pasando por Descartes y Leibniz (cf. Pasnau (2007), p. 306).

<sup>32</sup> Tomamos prestada, aunque de otro contexto, esta ingeniosa expresión de Husserl (1993), p. 111.

<sup>33</sup> Platón, desilusionado con los primeros filósofos de la naturaleza, se obliga a emprender una “segunda navegación” (cf. Platón, *Fedón*, 96a-99d). En ese pasaje, con asombrosa explicitación, Sócrates, a punto de ser ejecutado en la cárcel, enseña a sus discípulos la imposibilidad de explicar su fatídica decisión mediante meras descripciones de su “anatomía”, diríamos hoy. Cf. también Reale y Antiseri (1995), pp. 127-128.

<sup>34</sup> Canguilhem (1980), pp. 81-98.

constituye el sentido, el significado. Precisamente porque “tenemos la llave del sentido”, podemos entrar o salir, abrir o cerrar. En el reino animal, somos los que conocen el *lógos* (ζῶον λογικόν), y lo utilizamos de infinitos modos, a veces bien, a veces mal.

Pero el breve texto de Canguilhem condensa otros tópicos ya vistos aquí: se nos vuelve a señalar como punto clave el factor *semántico*, factor del que “las explicaciones de tipo fiscalista no han logrado nunca dar cuenta” y, por tanto, que tipifican una imposibilidad *de principio – de iure, no de facto –*; se nos habla explícitamente de los vanos intentos de “localizar” nuestra capacidad de significación, de la necesidad de admitir un “usuario real”, diríamos, que por lo demás, es *libre* y puede “jugar” con el sentido, en fin, prácticamente retoma todo nuestro recorrido hasta ahora, reiterando muy convenientemente la importancia decisiva de la libertad.

Es así como arribamos ya al último ejemplo que nos parece a la vez actual e instructivo: John Searle, uno de los filósofos contemporáneos que más ha insistido en la irreductibilidad de nuestra dimensión semántica desde los años '80, sobre todo con la publicación de su famosa alegoría de la “habitación china” (*the chinese room*)<sup>35</sup>:

Si la computación fuera suficiente para la comprensión (*cognition*), entonces cualquier agente que careciera de una capacidad cognitiva podría adquirir esa capacidad simplemente implementando el programa informático apropiado para manifestar esa capacidad. Imaginemos un hablante nativo de inglés – yo, por ejemplo –, que no entiende nada de chino. Imaginemos que estoy encerrado en una habitación con cajas de símbolos chinos (la base de datos) junto con un libro de instrucciones en inglés para manipular los símbolos (el programa). Imaginemos que personas que están fuera de la habitación envían pequeños conjuntos de símbolos chinos (preguntas) y estos son la entrada (*input*). Lo único que sé es que estoy recibiendo conjuntos de símbolos que para mí no tienen ningún sentido. Imaginemos que sigo el programa que me instruye cómo manipular los símbolos. Imaginemos que los programadores que diseñaron el programa lo hicieron tan bien, y yo me vuelvo tan bueno manipulando los símbolos chinos, que soy capaz de dar respuestas correctas a las preguntas (el *output*). El programa me permite, en la habitación, pasar la *Prueba de Turing* para entender el chino, pero pese a todo, no entiendo ni una sola palabra de chino. El punto del argumento es que si yo no entiendo chino sobre la base de implementar el programa apropiado para entender chino, entonces ninguna computadora digital lo podrá hacer sobre esa misma base solamente, porque la computadora, en cuanto computadora, no tiene nada que yo no tenga<sup>36</sup>.

Pese a la claridad del argumento, desde su primera aparición no dejó de suscitar polémica, lo que también ha sido un estímulo invaluable para que muchos filósofos de la mente, neurocientíficos y especialistas en computación se enfocaran en la comprobación – o no – de toda su estructura argumentativa, así como de sus conclusiones. A nuestro entender, la cuestión semántica es la más importante de ellas y la que logra resistir a todas las objeciones planteadas. Así la presenta el mismo Searle en uno de sus últimos textos al respecto: “Sintaxis no es semántica. La sintaxis por sí misma no es constitutiva de una semántica ni es por sí misma suficiente para garantizar la presencia de semántica”<sup>37</sup>.

No podemos abordar en este trabajo ni siquiera mínimamente la problemática que se ha desencadenado a partir de esta analogía de Searle, que ya pasa a ser un clásico en la filosofía de la mente y de la informática<sup>38</sup>. De todos modos, el *quid* de la cuestión puede elucidarse si volvemos a poner el foco en el “enigma del sentido”<sup>39</sup>:

La clave [de mi argumento] es ésta: una computadora digital es un aparato que manipula símbolos, sin cualquier referencia a su significado o interpretación. Los seres humanos, por otro lado, cuando piensan, hacen mucho más que eso. Una mente humana tiene pensamientos significativos, sentimientos, y contenidos [mentales] en general. Los símbolos formales por ellos mismos nunca pueden ser suficientes para ser contenidos mentales, puesto que los símbolos, por definición, carecen de significados (o interpretación, o semántica), a no ser que alguien desde afuera del sistema se los dé<sup>40</sup>.

<sup>35</sup> Otro filósofo, posiblemente tan conocido como Searle – y que ha sostenido decididamente el carácter *sui generis* de la conciencia – es Thomas Nagel (1937). En su célebre artículo de 1974 (*What is it like to be a bat?*) subraya la imposibilidad de situarnos en el punto de vista único de cualquier otro organismo consciente, demostrando con múltiples argumentos la irreductibilidad del plano subjetivo al objetivo. En 2014, consolida su crítica al reduccionismo científico de corte materialista e invita a repensar los fundamentos de la ciencia de tal forma que pueda admitir tres realidades por ahora excluidas o deformadas: la conciencia, el conocimiento y el valor. Hemos preferido a John Searle simplemente porque sus argumentos se oponen de forma más directa y explícita a un reduccionismo específicamente informático.

<sup>36</sup> Searle (2009): [http://www.scholarpedia.org/article/Searle's\\_Chinese\\_room#Three\\_misinterpretations](http://www.scholarpedia.org/article/Searle's_Chinese_room#Three_misinterpretations) (última actualización: 22/11/2024).

<sup>37</sup> *Ibid.*

<sup>38</sup> Para un tratamiento prácticamente exhaustivo y actualizado, cf.: <https://plato.stanford.edu/entries/chinese-room/> (última actualización: 15/11/2024). La validez del argumento, estimamos, se mantiene independientemente del “naturalismo biológico” que el mismo Searle dice profesar.

<sup>39</sup> Tomamos prestada esta expresión de Desilet (2023). No hace falta decir que ese otorgamiento originario de sentido –de interés ya milenario en Occidente– fue puesto de relieve muy especialmente por Edmund Husserl (para no seguir con una larga lista de pensadores): “Las vivencias intencionales se presentan como *unidades* en virtud de la operación de dar sentido (*Sinngebung*)” (Husserl (1993), p. 203).

<sup>40</sup> Searle (1989), p. 45. Este es el caso no sólo para los símbolos formales, sino para cualquier unidad de sentido o, para ser más claro, lo que solemos llamar “una cosa” o “un objeto”, cualesquiera sean. Ya Platón nos había enseñado que la pura materia (*lo visible*,

La computación existe solo en relación con algún agente u observador que impone una interpretación computacional a algún fenómeno. Este es un punto obvio. Debería haberlo visto hace diez años, pero no lo hice<sup>41</sup>.

En efecto, ¿puede el fenómeno (o proceso) de la significación ser explicado o reducido a términos no semánticos? ¿No ha sido precisamente esa la imposibilidad con la que una y otra vez tropezamos en este estudio al verificar el fracaso de tantos intentos? Ya nos lo había dicho Brigitte Falkenburg: “no veo por ningún lado otros conceptos científicos que establezcan un *punte semántico* sólido entre los fenómenos físicos y mentales”<sup>42</sup>. Un texto más reciente, secundando a Searle, es todavía más explícito:

(...) no importa cuán inteligente se comporte una computadora: puesto que los símbolos que procesa no tienen sentido para ella, en realidad no es inteligente; no está pensando. Sus estados y procesos internos, siendo puramente sintácticos, carecen de semántica (o significado), por lo cual la computadora realmente no tiene estados mentales intencionales (es decir, significativos). (...) Las operaciones realizadas por la computadora son “formales” en el sentido de que responden solamente a la forma física de una secuencia de símbolos. Las mentes, al contrario, tienen estados con significado, con contenidos mentales. Nosotros atribuimos significado a las palabras y a los signos lingüísticos, y sólo respondemos a esos signos debido a que tienen significado [para nosotros], no por su mera apariencia física. En definitiva: entendemos. (...) Así, aunque las computadoras son capaces de manipular sintaxis para producir respuestas apropiadas a entradas (*input*) de lenguaje natural, no entienden las respuestas que arrojan (*output*), porque no pueden asociar significado a las palabras<sup>43</sup>.

Es así como se nos va haciendo cada vez más difícil concebir aquel “puente semántico” reclamado por Falkenburg, y cada vez más evidente “un verdadero abismo de sentido”<sup>44</sup> que confirma reiteradamente la inconmensurabilidad entre los dominios en cuestión.

Pese a tan sólidos argumentos, sin embargo, ¿cómo alejar la “tentación” de los distintos reduccionismos si ni siquiera podemos dar una definición cabal de la *significación*? ¿No estará también ella “mitificada”<sup>45</sup> y, por lo tanto, igualmente destinada a ser reemplazada por una racionalidad más empirista y más técnica?

Volvamos a insistir, una vez más, en el carácter *originario* que reviste el *sentido*. Dale Jacquette, citando a Aristóteles, muestra con toda claridad cómo un modelo teórico consistente (*i.e.*, que no derive finalmente en circularidad lógica ni en un regreso *ad infinitum*) se apoya a fin de cuentas en términos “primitivos” (*primitives*), imprescindibles e irreductibles a elementos más simples. Jacquette denomina a ese marco referencial teórico “modelo fundacional” (*foundational model*):

Toda teoría digna de ese nombre debe incluir conceptos básicos, primitivos, atómicos o indefinibles que se utilizan como “ladrillos” en la construcción de conceptos más complejos. El análisis filosófico, como el análisis químico, deberá en algún momento descomponer su tema de estudio en elementos irreductibles. Esto sólo puede ocurrir si hay conceptos indefinibles en base a los cuales se definen los demás. / En este modelo, toda teoría tiene [conceptos] primitivos; sólo es cuestión de qué primitivos aceptar y cómo se van a utilizar como fundamentos de la estructura global de una teoría. La teoría de la intencionalidad da carta de ciudadanía a las conexiones intencionales entre el pensamiento y sus objetos como conceptualmente primitivas o no pasibles de un análisis posterior, y de esta manera reconoce la intencionalidad como el fundamento de las definiciones y explicaciones filosóficas<sup>46</sup>.

Era lo que Edmund Husserl decía en sus *Investigaciones Lógicas*:

Lo que sea “significación” (*Bedeutung*) es cosa que puede sernos tan inmediatamente dado como lo que sea color y sonido. No se puede definir más detalladamente. Es un término descriptivamente último (*ein deskriptiv Letztes*). Cuando emitimos o comprendemos una expresión, esta expresión significa algo para nosotros, tenemos conciencia actual de su sentido. Este comprender, significar, emitir un sentido (*Sinn Vollziehen*) no es oír los sonidos verbales o vivir alguna imagen simultánea. Y así como nos son dadas diferencias fenomenológicas evidentes entre los sonidos, así también nos son dadas diferencias entre las significaciones<sup>47</sup>.

---

léase: todo lo que puedan abarcar nuestros sentidos, incluso ampliados por la técnica) no es en absoluto suficiente como para hacer surgir una simple *unidad de sentido*, como por ejemplo, un naranjo (cf. Platón, *Fedón*, 65d; 4d-75c). Esa *unidad de sentido* es prerrogativa exclusiva de aquella capacidad enigmática de otorgar sentido (*Sinngebung*). Sin embargo, dejamos abierta la posibilidad de que la *constitución* que entraña se deba en último análisis a una génesis absoluta desde el *Ego Transcendental*, o más bien se trate de un *reconocimiento* de sentido, según una tradición más aristotélica.

<sup>41</sup> Searle (2002), p. 17 (publicado por primera vez en 1993).

<sup>42</sup> Falkenburg, *op. cit.*, p. 365.

<sup>43</sup> Gennaro (2016), pp. 170-171. Roger Schank, un reconocido especialista en educación e IA nos dice: “La computadora puede inventar una historia muy ocurente para contarnos, y elegir la ocasión más oportuna. Pero, por supuesto, no sabe lo que está diciendo” (Schank (2015), p. 133).

<sup>44</sup> Para utilizar nuevamente una expresión husserliana (Husserl, *op.cit.*, 114).

<sup>45</sup> Cf. Kripke (1982), p. 51.

<sup>46</sup> Jacquette (1989), pp. 620-621.

<sup>47</sup> Husserl (2006), p. 352. Cotejamos la traducción con el original: Husserl (2009), p. 187; cf. Ogden (1985), p. 272. Así, no hay que esperar a Michel Foucault para hacer *arqueología*, si por ello entendemos la búsqueda del sentido más originario. De hecho, según Eugen Fink – el conocido asesor de Husserl – respecto del término “arqueología”, “Husserl siempre se lamentó de que una expresión que capta verdaderamente la esencia de la filosofía haya sido ya acaparada por una ciencia positiva” (Fink (1966), p. 199; cf. Geraets (1971), p. 142).

Este es el caso de la Fenomenología, pero también es el caso de la Química y de la Física (piénsese en las nociones “primitivas” u originarias de “fuerza” y de “causa-efecto”, etc.). A este tenor cobra todo su sentido aquella frase de Wittgenstein: “Si lo verdadero es lo fundado, entonces el fundamento no es verdadero, pero tampoco falso”<sup>48</sup>.

Creemos que en este punto la discusión y la reflexión se pueden estancar o, peor aún, descarrilar precisamente por no tener en cuenta aquello de lo que ya nos había alertado un filósofo francés, algo postergado en la historia de la Filosofía, Maine de Biran (1766-1824):

Comoquiera que se explique cómo el influjo cerebral pone en juego los nervios y por ellos los músculos, eso no hará que se conciba mejor, o tal vez sólo servirá para oscurecer y desnaturalizar la operación y la eficacia del querer en el movimiento voluntario. La conciencia tiene sus propios signos, incommunicables e intraducibles. El signo propio del movimiento voluntario es el esfuerzo que quiero y percibo a la vez en su libre determinación y en su resultado. No podemos ir más lejos o buscar explicaciones sin equiparar y confundir dos órdenes de hechos de índole totalmente distintas. (...) Cuando se trata de un hecho originario (*primitif*) del sentido íntimo, el pedir ver, imaginar, figurarse el cómo del hecho es no saber lo que se pregunta, preguntando lo que ya se sabe<sup>49</sup>.

¿Qué es lo que “ya se sabe”? ¿Qué, lo que todos ya sabemos? Si tomamos el verbo “saber” en su sentido lato (como lo toma Maine de Biran en el texto), y que abarca también el sentido de “experiencia originaria” o *fait primitif*, antes de saber cualquier cosa, somos primeramente conscientes de lo que nos inquieta<sup>50</sup> y nos hace buscar, preguntar, reflexionar. Cada uno de nosotros es el mayor experto en *significación* que se nos pueda presentar. Y ¿qué significamos originariamente? La “materia inerte”, nos contestaría Biran, a la que tenemos acceso a partir del momento en que se completa el circuito de la sensación (“circuito” enteramente explicable por la neurofisiología). Sólo a partir de ese anillo sensible que “rodea” y limita nuestra capacidad de significación, ésta tendrá algo a que tender (*in-tendere*). Pero, entonces Leibniz tenía razón:

Se me objetará el axioma admitido por los filósofos, según el cual “nada hay en el alma que no venga de los sentidos”. Pero hay que exceptuar al alma misma, y a sus afecciones. “*Nihil est in intellectu quod non fuerit in sensu*”, excipe: “*nisi ipse intellectus*”<sup>51</sup>.

Una y otra vez se nos depara, y prácticamente se nos impone a los ojos, este como co-principio de nuestro conocimiento en general: juntamente con aquello que nos ofrecen nuestros sentidos – cada vez más extendidos por la técnica – constatamos también esa simple capacidad de pensar, de significar, de *in-tendere* como igualmente necesaria e igual de interdependiente. Aquí cobra pleno sentido hablar de una *interioridad* y su correlativa *exterioridad*. La pura sensación no alcanza para explicar la génesis de la semántica; tampoco la intencionalidad de la conciencia sola, ejecutándose en el vacío<sup>52</sup>. Ambos principios se necesitan mutuamente para el nacimiento del significado, e incluso para el “nacimiento del yo”<sup>53</sup>.

Finalmente, si queremos avanzar aún más y hacer comparecer el momento crucial en la articulación de ambas esferas, es instructivo visitar a Aristóteles y el rol decisivo que atribuye a la *fantasía* (φαντασία), esa enigmática capacidad que tenemos “por la cual decimos que una imagen (φάντασμα) llega a ser para nosotros”<sup>54</sup>. Dado su índole estrictamente originaria (según la acepción husserliana ya mencionada), todo lo que puede decir Aristóteles es que dichas imágenes primeras se dan, ocurren (γίγνεσθαι) por ocasión de los diversos tipos de sensación (αἴσθησις), aunque no *causados* por ellos. Por tanto, a partir de las sensaciones, que *no son imágenes*, la fantasía genera por vez primera imágenes para la conciencia. La fantasía es simplemente lo que hace que lo *absolutamente no consciente* (la materia, en la que se resuelve completamente la sensación) sea originariamente consciente y se haga “luz”<sup>55</sup> en la conciencia,

<sup>48</sup> Wittgenstein (2005), p. 210.

<sup>49</sup> Maine de Biran (1990b), p. 325.

<sup>50</sup> Ya sea la *inquietudo cordis* de San Agustín, ya sea el “sentirnos menesterosos”, insatisfechos, de Ortega (cf. Ortega y Gasset (1983), p. 546. Sea como fuere, dichas disposiciones fundamentales, ¿las podremos encontrar alguna vez como vivencias de una computadora? ¿Tiene sentido decir que sí, o “tal vez”?

<sup>51</sup> Leibniz (1983), p. 119.

<sup>52</sup> Cf. Sartre (1943), p. 27.

<sup>53</sup> Literalmente, el filósofo francés rehace la génesis de la autoconciencia como “manifestación primera y el nacimiento del yo (*l'enfantement du moi*), que nace para sí mismo al comenzar a conocerse” (Maine de Biran (1990a), p. 76).

<sup>54</sup> De Anima III, 428a2-3: Εἰ δὲ ἔστιν ἡ φαντασία καθ' ἣν λέγομεν φάντασμα τι ἡμῖν γίγνεσθαι. Hemos utilizado la traducción y comentarios de Childs (2016), y el texto original de Aristóteles *De anima* (1912). Creemos encontrar aquí la génesis de los *qualia* que, para Aristóteles y según nuestra interpretación, son precisamente esas imágenes originarias. Los φαντάσματα aristotélicos, claro está, no son sólo imágenes visuales, sino reproducciones fenoménicas de cada “sensible propio” (ἴδιον αἰσθητόν: Aristotle, *De Anima II*, 418a11), es decir, los estímulos sensibles específicos de cada uno de los 5 sentidos.

<sup>55</sup> “Con esta segunda [clase de] vida *nace la conciencia*, que viene a iluminar *las tinieblas del organismo* (...)”, siendo éste caracterizado por una “fuerza ciega que *se ignora a sí misma*” (Maine de Biran (1990a), pp. 07-08). Para una mayor precisión acerca de esa “secreta” articulación entre conciencia y materia, cf. Porto Nogueira (2018), pp. 45-60.

para seguir con la comparación aristotélica: “Puesto que la vista es el principal de los sentidos, el nombre «fantasía» (φαντασία) viene de luz (φῶς), porque sin luz no es posible ver”<sup>56</sup>.

## 7. Conclusión

Tal vez ayude volver a mirar brevemente el itinerario transitado. Hemos partido de una descripción en primera persona de lo que una tradición milenaria ha dado por llamar en Occidente “interioridad” u *homo interior*. A continuación, hicimos comparecer distintos y diversos alegatos, desde el campo filosófico, científico y computacional elaborados en orden a socavar los fundamentos y el sentido mismo de esa interioridad. Al examinar dichos cuestionamientos, nos encontramos con la tesis de la *inconmensurabilidad* de Brigitte Falkenburg, que señalaba la ausencia de un “puente semántico” entre lo físico y lo mental. Dicha inconmensurabilidad quedó muy bien respaldada por la incompatibilidad entre ambos órdenes de eventos. En este sentido, la metáfora del molino de Leibniz sólo vino a confirmarlo. Georges Canguilhem, por su parte, la expresa en términos de irreductibilidad de la relación que cada uno de nosotros tiene consigo mismo, con los demás, con las cosas, y la relación *entre* cosas, de la que esencialmente se distingue la primera. Y a tal punto somos los detentores del sentido que podemos perfectamente alterar esa “relación *con*”, subvirtiéndola y falseando la comunicación de los significados.

De ahí pasamos a la “habitación china” de John Searle. Entonces volvimos a constatar, una y otra vez, la irreductibilidad de la semántica a cadenas casi infinitas de signos, pero cuyos significados deben ser atribuidos “desde afuera del sistema”. Más adelante, la dicotomía se nos mostró como entre signos o secuencia de signos *físicos* – que sólo pueden ser procesados también físicamente – por un lado – , y la relación significativa, de otorgamiento de sentido (*Sinngebung*), por otro. Y a fin de que el “enigma de la significación” no quedara envuelto entre nubes de ambigüedades, Dale Jacquette, remitiéndose a Aristóteles, justifica la necesidad de términos primitivos “no pasibles de un análisis posterior”.

Para seguir con Aristóteles, llevamos nuestra “arqueología” hasta los límites de lo expresable al poner en relieve el rol decisivo de la *fantasía* – o como se quisiera llamar la facultad mental que haga surgir las primeras imágenes (originarias, por tanto) en la conciencia – . Esas imágenes primordiales, pese a estar todavía desprovista de significación, son empero necesarias para detener de algún modo el impetuoso y multiforme río heracliteano de nuestras sensaciones que, traducidas a imágenes por la fantasía, ya pueden quedar disponibles en la memoria, para finalmente permitir a nuestra inteligencia, dadora de sentido, realizar lo que le es propio – todo esto según el esquema aristotélico, claro está – .

Sin embargo, a pesar de que entre ese proceso semántico y la materia “ignorante de sí misma” se abra aquel “verdadero abismo de sentido”, no podemos soslayar el hecho de que estén completamente integrados en el ser humano. En este estado de cosas, nace una propuesta novedosa por parte de Maine de Biran y de otro de los autores ya citados: David Chalmers. Más allá de reconocer el carácter *sui generis* de la conciencia y de nuestra capacidad de significación, estos filósofos nos proponen una especie de “ciencia mixta”<sup>57</sup> o integradora que se tome en serio ambas competencias sin confundirlas ni derivarlas una de otra. El primero de ellos, se dirigía así a la Sociedad Médica de su tiempo:

El asunto particular que propongo hoy a su consideración se encuentra, por así decirlo, en los confines de dos ciencias que abarcan todo el hombre; pertenece igualmente a la Fisiología – que considera ese ser mixto como siendo simplemente un ser vivo y sensible – ; y pertenece a la Psicología – que lo considera, además, como inteligente y pensante<sup>58</sup> .

En cuanto a la investigación llevada a cabo por Chalmers, al menos es plausible esperar para el futuro próximo una ciencia consistente de la conciencia, que “tome seriamente los *data* obtenidos en primera persona de tal modo que se pueda estudiar una articulación entre esos datos y los arrojados en tercera persona sin caerse en un reduccionismo<sup>59</sup>” en ninguna de las dos direcciones.

Para terminar, reconozcamos que el reduccionismo naturalista o el computacional no son las únicas simplificaciones del “fenómeno humano”. También la Fenomenología ha de resistir a una tentación, pero de signo opuesto: la de reducir la materia y el cuerpo biológico a mera *significación* o a una cadena infinita de interpretaciones. Las corrientes más actuales de la Fenomenología parecen estar incluso todavía más propensas a esa especie de *monismo fenomenológico*<sup>60</sup>, sobre cuya base se hace muy difícil la posibilidad

<sup>56</sup> Aristotle, *De Anima*, 430a3ss. Así, la solución para el “problema difícil” que nos plantea la existencia de la conciencia (¿por qué hay actividad consciente paralelamente a la actividad cerebral, o por qué hay conciencia en absoluto?) no está en buscar qué características neurofisiológicas explican o exigen el surgimiento de la conciencia (cf. Northoff (2014), pp. 374-377), sino en cómo el sujeto del conocimiento, en su primer acto consciente, se apropia (hace suyas) las sensaciones más básicas, dando así lugar a toda la inagotabilidad de la vida subjetiva.

<sup>57</sup> Azouvi (1995), pp. 139ss.

<sup>58</sup> Maine de Biran (1984), p. 10.

<sup>59</sup> Chalmers (2010), pp. 41ss.

<sup>60</sup> Esta postrera desviación de la Fenomenología ha sido claramente señalada por Emmanuel Falque (1963 - ), aunque no en esos términos. Falque va al rescate de un “verdadero cuerpo” en contraste con la *chair* fenomenologizada: “No simplemente la vivencia interna del cuerpo que se auto-afecta (como lo interpreta erróneamente Michel Henry en ‘*Incarnation*’), sino un verdadero cuerpo,

de diálogo, no solo con las ciencias y el sentido común, sino también con una importante parcela de la filosofía contemporánea.

## 8. Referencias bibliográficas

- Aristotle (2016): *De Anima*, Oxford, Clarendon Press.
- Azouvi, F. (1995) : *Maine de Biran: la science de l'homme*, Paris, Vrin.
- Bech, J.M. (2005): *Merleau-Ponty. Una aproximación a su pensamiento*, Barcelona, Anthropos.
- Bengio, Y. et. al. (2021): "Deep Learning for AI", *ACM*, 64, pp. 58-65.
- Bennett, M.; Hacker, P. (2003): *Philosophical Foundations of Neuroscience*, Malden, Blackwell.
- Canguilhem, G. (1980) : "Le cerveau et la pensée", *Prospective et Santé*, 14, pp. 81-98.
- Carney, J. D., et al. (1964): *Fundamentals of Logic*, New York, Macmillan.
- Castro, E. (2011): *Diccionario Foucault, Temas, conceptos y autores*, Buenos Aires, Siglo Veintiuno.
- Chalmers, D.J. (2010): *The Character of Consciousness*, New York, Oxford University Press.
- Chomsky, N.; Foucault, M. (2006): *The Chomsky-Foucault debate*, New York, The New Press.
- Cole, D. (2024): "The Chinese Room Argument" en Zalta, E.N.; Nodelman, U. (Eds.), *The Stanford Encyclopedia of Philosophy*, (última actualización: 15/11/2024), <https://plato.stanford.edu/archives/win2024/entries/chinese-room/>
- Costa, J. P. (2022) : "Un tournant? Interview d'Emmanuel Falque", *Revista Filosófica de Coimbra*, 31, pp. 279-290.
- Dennett, D. (1995): *La conciencia explicada*, Barcelona, Paidós.
- Falkenburg, B. (2019): *Mythos Determinismus: Wieviel erklärt uns die Hirnforschung?*, Berlin, Springer.
- Fink, E. (1966): *Studien zur Phänomenologie 1930-1939*, La Haye, Martinus Nijhoff.
- Foucault, M. (2004): *El pensamiento del afuera*, Valencia, Pretextos.
- Frankish, K. (ed.) (2017): *Illusionism as a theory of consciousness*, Exeter, Imprint Academic.
- Gabriel, M. (2019): *I am Not a Brain*, Cambridge, Policy Press.
- Gabriel, M. (2024): *Sense, Nonsense, and Subjectivity*, Cambridge M., Harvard University Press.
- Gennaro, R.J. (2016): *Consciousness*, New York, Taylor & Francis.
- Geraets, T. F. (1971) : *Vers une nouvelle philosophie transcendantale. La genèse de la philosophie de Maurice Merleau-Ponty jusqu'à la Phénoménologie de la perception*, La Haye, Martinus Nijhoff.
- Goldstein, E. B. (2020): *The Mind: Consciousness, Prediction, and the Brain*, Cambridge-Massachusetts, MIT Press.
- Gregory, G. (2023): *The Enigma of Meaning*, Jefferson N.C., McFarland & Company.
- Hamblin, C. L. (2016): *Falacias*, Lima, Palestra.
- Husserl, E. (1993): *Ideas relativas a una fenomenología pura y una filosofía fenomenológica*, Madrid, Fondo de cultura económica.
- Husserl, E. (2006): *Investigaciones lógicas*, Madrid, Alianza.
- Husserl, E. (2009): *Logische Untersuchungen II*, Hamburg, Felix Meiner.
- Husserl, E. (1996): *Meditaciones cartesianas*, México, Fondo de cultura económica.
- Jacquette, D. (1989): "Adventures in the Chinese Room", *Philosophy and Phenomenological Research*, 4, pp. 620-621.
- Krauss, P. (2024): *Artificial Intelligence and Brain Research*, Berlin, Springer.
- Kripke, S. A. (1982): *Wittgenstein on Rules and private language*, Cambridge MA, Harvard University Press.
- Leibniz, G. W. (1981): *Monadología*, Oviedo, Pentalfa.
- Leibniz, G.W. (1983): *Nuevos ensayos sobre el entendimiento humano*, Madrid, Nacional.
- Maine de Biran (1984) : *Œuvres*, t. V, Paris, Vrin.
- Maine de Biran (1990) : *Œuvres*, t. X-2, Paris, Vrin.
- Maine de Biran (1990) : *Œuvres*, t. XI-3, Paris, Vrin.
- Merleau-Ponty, M. (1945) : *Phénoménologie de la Perception*, Paris, Gallimard.
- Nagel, T. (1974) : "What is it like to be a bat?", *The Philosophical Review*, 83, pp. 435-450.
- Nagel, T. (2014): *La mente y el cosmos*, Madrid, Biblioteca Nueva.
- Northoff, G. (2014): *Minding the Brain*, New York, Palgrave MacMillan.
- Ogden, C. K.; Richards, I. A. (1985): *The meaning of meaning*, New York, Ark.
- Ortega y Gasset, J. (1983): *Obras completas, tomo IV*, Madrid, Alianza.
- Pasnau, R. (2007): "Mind and Extension (Descartes, Hobbes, Moore)" en Lagerlund, H. *Forming the Mind*, Dordrecht, Springer, pp. 283-310.
- Platón (1988): *Diálogos III. Fedón, Banquete, Fedro*, Madrid, Gredos.
- Porto Nogueira, T. (2016): *¿Una fenomenología realista de la percepción?*, Córdoba-Argentina, EDUCC.

---

'en carne y hueso' – *Körper* (cuerpo material y biológico) y no solamente *Leib* (vivencia del cuerpo), para quedarnos en los términos de la fenomenología (...)" (Costa (2022), p. 284).

- Porto Nogueira, T. (2018): "Filosofía y Ciencia: aportes para un mayor entendimiento" en De Asúa, M.; Figueroa, M. (ed.), *Ciencia, Filosofía y Religión*, Buenos Aires, Ágape.
- Reale, G.; Antiseri, D. (1995): *Historia del pensamiento filosófico y científico, tomo I*, Barcelona, Herder.
- Sant'Agostino (1930): *De Magistro. De vera religione*, Firenze, Vallecchi.
- Sartre, J-P. (1943): *L'Être et le néant*, Paris, Gallimard.
- Schank, R. (2015): "Machines that Think are in the Movies", en Brockman, J. (ed.), *What to Think About Machines that Think*, New York, Harper Collins, pp. 132-135.
- Searle, J. (1989): "Artificial Intelligence and the Chinese Room: An Exchange", *The New York Review*, 36, pp. 44-45.
- Searle, J. (2009): "Chinese room argument", *Scholarpedia*, 4 (última actualización: 22/11/2024), [http://www.scholarpedia.org/article/Searle's\\_Chinese\\_room#Three\\_misinterpretations](http://www.scholarpedia.org/article/Searle's_Chinese_room#Three_misinterpretations).
- Searle, J. (2002): *Consciousness and Language*, Cambridge, Cambridge University Press.
- Stein, E (1996). *Ser finito y Ser Eterno*, México, Fondo de cultura económica
- Uña Juárez, A. (2001): "San Agustín: interioridad, reflexividad y certeza", *Revista Española de Filosofía Medieval*, 8, pp. 31-52.
- Wittgenstein, L. (2005): *On Certainty*, Oxford, Blackwell.