

# EL CONOCIMIENTO DE LA MENTE: PARADOJAS DE UN RÉGIMEN PREPOSICIONAL REVERSIBLE

MIQUEL CANDEL  
Universitat de Barcelona

## RESUMEN

El artículo pasa sucintamente revista a las principales teorías sobre la naturaleza de la mente, desde la Antigüedad clásica hasta nuestros días, y toma partido por una concepción monista materialista que, sin embargo, preserva la especificidad de la mente (el *yo*) como nudo relacional y centro irreductible de perspectiva que, precisamente por eso, no puede nunca adquirir una visión objetiva de sí mismo como puro sujeto.

## ABSTRACT

This paper succinctly reviews the main theories concerning the nature of mind, from the Antiquity to the present day, and adopts the view of a kind of materialistic Monism, which, this notwithstanding, preserves the specific character of mind (the *self*) as a relational node and an irreducible centre of perspective. This is the very reason for the fact that it can never get an objective vision of itself as a pure subject.

## Del mundo de las sombras al mundo de las formas

Desde que Aristóteles escribiera su célebre tratado *Perì psychês* o *De anima*, la mente se ha convertido en uno de los temas filosóficos por antonomasia. Y no sólo por la inercia de los comentaristas antiguos, medievales e incluso modernos, con su machacona y reiterada (aunque no siempre reiterativa) exégesis de ese revolucionario texto. Motivos muy distintos de los que inspiraron a Aristóteles movieron a los estoicos y a pensadores cristianos seminales, como Agustín de Hipona, a buscar en el «interior» la clave de la sabiduría (es decir, el objeto por excelencia de la filosofía): la seguridad que ofrece esa «ciudadela interior»<sup>1</sup> frente

---

1. Véase la obra homónima de Pierre Hadot: *La citadelle intérieure*, París, Fayard, 1992. El diálogo *Alcibíades I*, atribuido a Platón, anticipa ya esta orientación de la filosofía hacia el autoconocimiento.

a la tornadiza suerte que condiciona cualquier proyecto vital hace que la mirada introspectiva pase a ocupar lugar preferente entre los esfuerzos del ser humano amante del saber.

Pero no se puede decir que esa «interioridad» sea un punto de partida, el hogar familiar del ser humano, al que periódicamente regresa a reposar de su fatigosa «lucha por la vida». Es, en todo caso, un hogar que el ser humano se ha ido construyendo lentamente a partir de una visión casi totalmente opuesta, según la cual el mundo circundante estaba animado desde dentro por un (conjunto de) «espíritu(s)» mientras el ser humano era apenas una marioneta cuyos miembros se hallaban temporal y extrínsecamente unidos por un soplo vital precario hasta que otro soplo, el de la «negra muerte», los separaba.<sup>2</sup> En esa primera etapa de lo que Bruno Snell llamó «el descubrimiento del espíritu»,<sup>3</sup> el «alma» (*psychē*) no era la intimidad constitutiva del yo, sino un vínculo extrínseco entre las partes del cuerpo y, tras la muerte, una lánguida sombra que se arrastraba a duras penas por los desolados páramos del Averno.

Tuvo que sobrevenir un drástico cambio social en lugares como Atenas (pero ya antes en los emporios mercantiles de Jonia y la Magna Grecia) para que el alma dejara de ser un simple hálito,<sup>4</sup> es decir, una corporeidad vaga y blanda (*animula vagula blandula*, la llamará el emperador Adriano) y se convirtiera en una «singularidad» para la que había que habilitar nuevos espacios conceptuales (y ontológicos). El mencionado cambio social fue el ascenso de clases no vinculadas directamente a la economía agraria, con la consiguiente introducción, en la visión de la realidad, de paradigmas «formales» que ponen en primer plano la *relación* y el *orden* entre los elementos constitutivos de las cosas y relegan al trasfondo los elementos mismos. De ahí teorías como la citada por Aristóteles en *De anima* I 4, según la cual el alma sería la «armonía» de los elementos integrantes del cuerpo. El propio Aristóteles, recogiendo el testigo de Platón, acabará calificando el alma de:

- 
2. Como medita el personaje de Julio César en la homónima ópera de Händel:

*Misera vita,  
O quanto è frat tuo stato!  
Ti forma un soffio  
E ti distrugge un fiato.  
(Acto I, escena III)*

3. *Die Entdeckung des Geistes*, Hamburgo, 1946 (*Las fuentes del pensamiento europeo*, Madrid, Razón y Fe, 1965).

4. Ésa, al fin y al cabo, es la etimología de *psychē*. Con posterioridad a Aristóteles, los estoicos recuperaron (en lo que no hay más remedio que considerar como un retroceso intelectual) el concepto de «alma-aliento» con su noción de *pneûma*, tradicionalmente vertido al homólogo latino de 'hálito': *spiritus* (de donde nuestro 'espíritu').

## Lugar de las formas

«Dicen bien los que dicen que el alma es el lugar de las formas, si exceptuamos que no lo es toda ella, sino sólo la intelectiva y que no es las formas en acto, sino en potencia.»<sup>5</sup> En efecto, tras una «disección» de las funciones del alma,<sup>6</sup> Aristóteles centra su análisis en la llamada facultad intelectiva, o *noûs*, que la tradición llamaría en latín *intellectus* o *mens*, *intelecto* (*entendimiento*) o *mente*. Y desarrollando la concepción anaxagórea de la mente<sup>7</sup> al hilo de su teoría hilemórfica, caracteriza el *noûs* como un reverso inmaterial de la materia primera, con la que comparte, paradójicamente, la total carencia de forma propia:

Por lo que se refiere a aquella parte del alma con que el alma conoce y piensa –ya se trate de algo separable, ya se trate de algo no separable en cuanto a la magnitud, pero sí en cuanto a la definición– ha de examinarse cuál es su característica diferencial y cómo se lleva a cabo la actividad de inteligir. Ahora bien, si el inteligir constituye una operación semejante a la sensación, consistirá en padecer cierto influjo bajo la acción de lo inteligible o bien en algún otro proceso similar. Por consiguiente, el intelecto –siendo impasible– ha de ser capaz de recibir la forma, es decir, ha de ser en potencia tal como la forma pero sin serlo ella misma, y será respecto de lo inteligible algo análogo a lo que es la facultad sensitiva respecto de lo sensible. Por consiguiente y puesto que entiende todas las cosas, necesariamente ha de ser sin mezcla –como dice Anaxágoras– para que pueda dominar o, lo que es lo mismo, conocer, ya que lo que exhibe su propia forma obstaculiza e interfiere a la ajena. Luego no tiene naturaleza alguna propia aparte de su misma potencialidad. Así pues, el denominado intelecto del alma –me refiero al intelecto con que el alma razona y enjuicia– no es en acto ninguno de los entes antes de inteligir.<sup>8</sup>

Este texto inaugura un enfoque muy particular del estudio de la mente. Enfoque que, sobre una base que podríamos llamar «materialista» en la concepción de la relación alma-cuerpo, admite la peculiaridad del intelecto como aquella función del alma que carece de vínculo esencial con el cuerpo (aunque no necesariamente de vínculo espacio-temporal: «no separable en cuanto a la magnitud [*mégethos*], pero sí en cuanto a la definición [*lógon*]») y que, a dife-

5. *De anima* III 4, 429a27-29.

6. El término 'disección' propone una metáfora muy impropia para reflejar el sentido del análisis aristotélico, ya que la distinción entre las funciones anímicas no es la que se da entre las partes coordinables de un todo (partes que se hallan, por así decir, en el mismo plano), sino la existente entre los grados de una escala, donde los grados superiores contienen los inferiores.

7. Frs. DK 12 B y 14 B.

8. *De anima* III 4, 429a10-24.

rencia de los sentidos, no sólo no se deja impresionar de manera duradera por la forma del objeto conocido, sino que carece por sí mismo de cualquier clase de forma; en efecto, de tener o retener alguna forma, ésta podría obstaculizar, como ocurre con los sentidos, la percepción de otras formas diferentes.

En otras palabras, la capacidad de conocer (mediante) las formas, objeto (o, mejor, *medio*) de conocimiento por antonomasia,<sup>9</sup> presupone una absoluta versatilidad en el propio ser de la facultad cognoscente, que la distingue radicalmente de cualquier ente material (pero la coloca al mismo nivel de absoluta potencialidad que la materia). Podíamos decir, pues, que para Aristóteles la mente es inmaterial a fuerza de «ultramaterial». Porque, por otro lado, sin perder en modo alguno su potencialidad, se identifica en cada caso (en cada acto de conocer) con el objeto conocido.

Pero la peculiaridad ontológica del *noûs* no se reduce a su identificación con el «objeto» inteligible, facilitada, en último término, como veíamos en la nota anterior, por el carácter «instrumental» del objeto. Esa «transparencia» que lo caracteriza lo hace identificarse en definitiva consigo mismo. Es decir, por ser inmaterial, «el intelecto es inteligible exactamente como lo son sus objetos; en efecto, tratándose de seres inmateriales, lo que entiende y lo inteligido se identifican toda vez que el conocimiento teórico y su objeto son idénticos».<sup>10</sup>

Seguramente no hay que entender esto como una formulación anticipada de la «autoconciencia» fichteana, sino como la de una –para nosotros, modernos– paradójica «heteroconciencia» radical que lleva al entendimiento, plenamente identificado con su objeto, a reconocerse plenamente en él. Sólo la esencia del Motor Inmóvil cosmológico, definida como «intelección de la intelección» (*noësis noëseos*), parece ser propiamente autoconciencia en cuanto que su objeto es totalmente interno e idéntico a su acto de conocer.

En cualquier caso, a partir de este atisbo de interiorización o repliegue de la mente sobre sí misma, la tradición platónico-aristotélica convergerá con la estoica para dar origen al llamado neoplatonismo, que tanto en sus variantes

9. En efecto, tal como se desprende de la noción misma de abstracción (*aphaíresis*) que maneja Aristóteles, las formas universales no son tanto aquello *que* la mente contempla *directamente* como aquello *a través de lo cual* o *en virtud de lo cual* la mente hace suyos los contenidos singulares que le suministran los sentidos. Esto ya es así en el mismo Platón, por más que la visión tópica de la teoría de las ideas (visión a cuya propagación contribuyó no poco el propio Aristóteles) las presente como contenidos epistémicos de superior grado de pureza y claridad, pero tan singulares como los contenidos de la percepción sensorial, es decir, como *paradigmas* (véase: Candel, M., *El nacimiento de la eternidad*, Barcelona, Ideas Books, 2002, cap. III). El carácter de objeto sólo lo pueden adquirir las llamadas «formas inteligibles» en tanto que *abstracciones*, es decir, como objetos indirectos o de *segundo orden*.

10. *De anima* III 4, 430a2-5.

pagana como cristiana (particularmente, la agustiniana) acabará substancializando el *noûs* como segunda *hypóstasis* del absoluto (o segunda persona de la Trinidad), despliegue de su unidad inefable a la vez que eslabón del hombre con la divinidad, a la que éste conocerá conociéndose a sí mismo.

## Del sujeto de atributos al sujeto de representaciones

Las filosofías musulmana y cristiana de la Edad Media proseguirán la tradición neoplatónica en su concepción del intelecto como doble vía de descenso y ascenso entre Dios y el hombre, combinada y complicada con la asociación, al menos en el caso musulmán, de las diversas esferas celestes con otros tantos intelectos (esa función la cumplen, en el cristianismo, las llamadas «jerarquías angélicas o celestiales»). Común a casi todas esas concepciones es la convicción de que el entendimiento, aun cuando se halle individualizado en múltiples seres inteligentes, es en el fondo uno, como unas son para todos las formas inteligibles (*universales*).

El nominalismo ockhamiano romperá esa tendencia al postular el carácter singular de todo contenido de conocimiento, incluido el intelectual, y la naturaleza meramente «objetiva» de los llamados «universales». <sup>11</sup> Como es natural, la singularidad de lo conocido se «contagia» al cognoscente: el *noûs* universal acaba estallando, al llegar la Edad Moderna, en una miríada de intelectos individuales, centros independientes de percepción y apercepción, transparentes a sí mismos pero sin comunicación directa alguna entre ellos. Las mónadas leibnizianas constituyen el ejemplo más acabado de la nueva concepción de la mente. Ellas suministran el paradigma de lo que será la mente moderna, centro de perspectiva y núcleo existencial por antonomasia; razón por la que acabará apropiándose en exclusiva del carácter de *sujeto*, que hasta entonces era común a toda realidad subsistente por sí misma, y recibirá el característico nombre de *yo*.

Suele decirse que es en Descartes donde se opera la transformación del pensar en sujeto, previa *substancialización* de su actividad, tal como parece indicar la definición de la mente como *res cogitans*. Semejante tesis, ampliamente admitida, se basa en una interpretación literal, «cosista», del término

---

11. Entiéndase por «objetivo», al menos en el uso ockhamista, aquello que no tiene más realidad que la de ser objeto de la mente, que sólo existe en el seno de ésta (acepción que es, prácticamente, la antítesis del uso actual del término). Por contra, 'subjetivo', en ese mismo uso ockhamista, es todo aquello que subsiste independientemente (y, en primer lugar, independientemente de nuestro conocimiento). Es ése el sentido original de 'sujeto' (*hypokeímenon*) en Aristóteles, que identifica en último término *sujeto* con *existencia* propiamente dicha: existir = subsistir (cf. Candel, M., *op. cit.*, cap. VII).

'res'. Ahora bien, dado que el contenido que Descartes asigna al acto *cogitans* es polifacético, constando de elementos que la tradición consideraba propios de «facultades del alma» distintas e irreductibles (desde la sensación hasta la volición, pasando por la imaginación y por todos y cada uno de los fenómenos que hoy consideramos patrimonio de la vida consciente),<sup>12</sup> parece más correcto sostener que Descartes concibe el *yo* del pensamiento, no como una substancia, sino como una actividad pluriforme aunque unifocal,<sup>13</sup> y que el uso de 'res' obedece simplemente a necesidades lógico-gramaticales, a saber: la imposibilidad de indicar la sustantivación de la forma verbal '*cogitans*' mediante el sencillo expediente de anteponer un artículo (disponible en casi todas las lenguas modernas pero inexistente en latín).

En Kant, finalmente, aparecen unidas de manera explícita las nociones de *yo* como centro de perspectiva incomunicable y de *sujeto* como base de lo existente por sí mismo. Véase al respecto la exposición del primer paralogismo de la razón pura:

*Substancia* es aquello cuya representación constituye el *sujeto absoluto* de nuestros juicios, aquello que no puede, por tanto, ser empleado como determinación de otra cosa. Yo, en cuanto ser pensante, soy el *sujeto absoluto* de todos mis juicios posibles, pero esta representación de mí mismo no puede ser usada como predicado de otra cosa.<sup>14</sup>

Claro que este razonamiento de Kant tiene, como paralogismo que es, una conclusión rechazada por el propio Kant, a saber: que el *yo* (o, en terminología tradicional, el *alma*) es *substancia*. Pero eso no obsta para que tengamos consumada, a partir de este planteamiento, la «revolución copernicana»: el substrato último de lo real (que es a lo que corresponde el uso filosófico del término 'sujeto') pasa a ser el *yo*, cuyos *objetos* de conocimiento ocupan el mismo lugar respecto de él que las *propiedades* respecto de la *substancia*. El idealismo posterior no hará sino llevar el proceso hasta sus últimas consecuencias: del mismo modo que las propiedades son aspectos constitutivos de la substancia, por lo que están, de hecho, incluidas en ella, así también los objetos de conocimiento quedan absorbidos en la esfera del *yo*, como momentos de su actividad autoconsciente. El esquema ontológico plotiniano se ha inver-

---

12. «Res cogitans. Quid est hoc? Nempe dubitans, intelligens, affirmans, negans, volens, nolens, imaginans quoque, et sentiens.» (*Meditaciones metafísicas*, meditación segunda, [9]).

13. A favor de esta interpretación no substancialista parecen hablar pasajes de la meditación segunda como aquél en que sostiene: «existo mientras pienso» (*existo ... quandiu cogito*, [7]), relativización temporal incompatible con la categoría de substancia, al menos tal como se entendía desde Aristóteles.

14. Kant, I., *Crítica de la razón pura*, libro II de la dialéctica trascendental, primer paralogismo.

tido por completo y la moderna versión del *noûs*, en lugar de aparecer como «exhibidor» que despliega en forma discursiva y *ad extra* el contenido del inefable *Uno* absoluto, usurpa la posición central de éste y lo convierte en contenido de su propia actividad centrípeta.

Obviamente, en la medida en que ese sujeto, tras la crítica kantiana, ha dejado de ser substancia, carece del «espesor» ontológico que caracterizaba el viejo *noûs* (al menos el plotiniano, aunque dudosamente el aristotélico). Por ello no es casual que hayamos dejado de referirnos a él con un nombre para designarlo con un simple pronombre: *yo*, cuya referencia cambia constantemente en función del reparto de papeles en el discurso. Ésa es la precaria «morada interior» que el ser humano empezó a construirse ya antes de abandonar las cavernas y que a duras penas remató la Ilustración.

Pero la filosofía del siglo XX, como reacción en gran parte al desmedido «interiorismo» psicologista de finales del XIX, empezó a derribar ese precario edificio cuando apenas su arquitecto había logrado cubrir aguas. Wittgenstein (al menos, en sus *Investigaciones filosóficas*) sentó las bases para reducir el *yo* del *yo pienso* al del *yo digo* y éste, en último término, al *se dice*. Y Gilbert Ryle, iniciador de la «escuela de Oxford», en paralelo a la evolución de la psicología empírica hacia el conductismo, ofició (con gran unción, eso sí, a diferencia de iconoclastas posteriores) el funeral por el *cogito* cartesiano en su decisiva obra *The Concept of Mind*,<sup>15</sup> verdadera carta fundacional del conductismo filosófico. Pero ese paso último hacia el definitivo vaciamiento ontológico del sujeto, lejos de ser simple estación término en el trayecto conceptual del *yo*, se convirtió en punto de partida de una profunda renovación de la psicología, tanto filosófica como empírica.

## La naturalización de la mente

Dice Ryle, al comienzo de la obra, ya clásica, antes citada:

Existe una doctrina acerca de la naturaleza y el lugar de la mente que tiene tanto predicamento entre los estudiosos, e incluso entre los profanos, que merece el calificativo de teoría oficial. [...] La doctrina oficial, que procede básicamente de Descartes, viene a decir más o menos lo siguiente: [...] todo ser humano posee un cuerpo y una mente. Algunos preferirían decir que todo ser humano es un cuerpo y una mente. Su cuerpo y su mente están normalmente combinados en un todo, pero después de la muerte del cuerpo su mente puede seguir existiendo y funcionando.

---

15. Londres, Hutchinson, 1949 (reedición en Penguin Books, 1990).

Los cuerpos humanos están en el espacio y se hallan sujetos a las leyes mecánicas que gobiernan todos los demás cuerpos en el espacio. Los procesos y estados corporales pueden ser examinados por observadores externos. Así, la vida corporal de un hombre es un asunto público en igual medida que las vidas de los animales y los reptiles e incluso que la existencia de los árboles, los cristales y los planetas.

Pero las mentes no están en el espacio ni sus operaciones se hallan sujetas a las leyes mecánicas. El funcionamiento de una mente no es observable por observadores ajenos; su existencia es privada. Sólo puedo tener conocimiento directo de los estados y procesos de mi propia mente. Cada persona, pues, vive dos historias paralelas: una que consiste en lo que le ocurre a su cuerpo, otra que consiste en lo que ocurre en su mente y con su mente. La primera de esas vidas es pública, la segunda, privada. Los acontecimientos de la primera historia son acontecimientos del mundo físico, los de la segunda son acontecimientos del mundo mental.

Ésta es, en síntesis, la teoría oficial. A menudo me referiré a ella, exagerando deliberadamente, como «el dogma del Espíritu [Ghost]<sup>16</sup> en la Máquina». Espero probar que es completamente falso, y no falso en cuestiones de detalle, sino de principio. No es simplemente una combinación de errores parciales. Es un gran error, error de un género especial. Es, en otras palabras, un error categorial [category-mistake]. Representa los hechos de la vida mental como si pertenecieran a un tipo o categoría lógica (o gama de tipos o categorías), cuando en realidad pertenecen a otro diferente. El dogma es, por consiguiente, un mito filosófico.<sup>17</sup>

¿Qué es lo que Ryle entiende por «error categorial». Él mismo da una serie de ejemplos como el siguiente: si un niño que contempla el paso de un desfile militar tras haber oído decir que lo que desfilará es la división número tal, al ver pasar antes sus ojos grupos de soldados y armas que la gente llama «compañías», «batallones», «baterías», etc., pregunta: «Y ¿cuándo va a pasar la división?», la criatura incurre precisamente en ese tipo de error que el autor llama (bastante aristotélicamente, por cierto) «categorial».

Es fácil deducir de ahí que Ryle entiende que la contraposición entre mente y cuerpo efectuada por los filósofos dualistas *à la* Descartes (y por la filosofía cristiana en general) es una burda confusión consistente en poner en el mismo plano ontológico dos aspectos formales diferentes de un mismo objeto material, con la consiguiente división de éste en dos objetos. Por desarrollar

---

16. El término inglés, *ghost*, sugiere peyorativamente la noción popular de espíritu como «fantasma», «espectro», etc.

17. *Ibid.*, pp. 13 y 15 de la edición de 1990.

un ejemplo más sencillo, aducido por Aristóteles al comienzo del *De anima*,<sup>18</sup> la confusión de la que hablamos sería aquella en que incurriría quien a la vista de un edificio dijera ver sólo un montón de ladrillos y preguntara dónde estaba la casa. La mente, pues, según Ryle, está con el cuerpo en una relación análoga a la existente entre, pongamos, el funcionamiento de un motor y el conjunto de las piezas que componen éste. El error de los dualistas consistiría en convertir en otras tantas sustancias dos aspectos formalmente diferentes de lo que no sería sino una sola y misma entidad material.

Como veremos hacia el final de este artículo, la crítica de Ryle es esencialmente atinada y se basa en una acertada comprensión del problema filosófico de fondo que plantea la reflexión sobre la mente. Como lo es también el enfoque del tema por Wittgenstein. Al menos en la medida en que rechaza los planteamientos tradicionales y, muy particularmente, los del psicologismo decimonónico. Pero lo cierto es que la terapia destinada a curar la «bulimia» psicologista ha acabado degenerando, a mi modo de ver, en una especie de «anorexia» reduccionista que seguramente permite decir una vez más aquello de que «ha sido peor el remedio que la enfermedad».

En primer lugar, el estricto conductismo filosófico nacido al socaire de la obra de Ryle aboca el estudio de la mente a una circularidad explicativa sin salida, como han mostrado desde finales de los años cincuenta sus críticos, Donald Davidson entre ellos.<sup>19</sup> Al pretender reducir toda descripción de los llamados «estados mentales» al análisis del comportamiento socialmente observable se incurre en un círculo vicioso hermenéutico entre, al menos, la caracterización de las *creencias* y la de los *deseos*. Así, por ejemplo, la supuesta *creencia* de Pedro en una inminente lluvia sólo se puede identificar con su acción de tomar el paraguas al salir de casa si se presupone (sin fundamento conductual alguno) que Pedro *desea* evitar mojarse; de no ser así, la acción de salir a la calle provisto de un paraguas podría identificarse con cualquier otra creencia (por ejemplo, que el paraguas puede resultar una buena arma defensiva en caso de un encuentro desagradable al atravesar ciertas calles donde suelen operar delincuentes). A la inversa, la interpretación de esa misma acción de Pedro como plasmación de su *deseo* de no mojarse sólo puede sostenerse en la previa asunción (inverificable dentro de este planteamiento) de que Pedro *crea* que va a llover.

---

18. *De an.* I 1, 403b1-10.

19. En el número 12 de sus *Essays on Actions and Events* (Oxford, Clarendon, 1980 (1982), p. 236), Davidson cuenta cómo, a raíz de su participación en un experimento destinado a estudiar ciertas condiciones del comportamiento racional humano, la incoherencia de los resultados obtenidos (si se los juzgaba con criterios conductistas) le impulsó a abandonar de una vez por todas esos planteamientos (y, con ellos, la práctica de la psicología experimental).

Nada de extraño, pues, que el conductismo estricto desapareciera poco a poco de la literatura dedicada al estudio de la mente. Desaparición que hay que circunscribir, sin embargo, a su forma canónica, tal como aquí se ha expuesto, pues en forma larvada o atenuada sigue presente en mayor o menor grado en casi todos los tratamientos naturalistas del problema de la mente (y, tal como veremos, con razón).

Al emplear la expresión «tratamiento naturalista» en el párrafo precedente hemos aludido al rasgo esencial de la filosofía contemporánea de la mente que permite distinguirla tanto de la antigua filosofía del *noûs* (que en su mayoría, desde Anaxágoras, contraponía mente y naturaleza) como de la moderna filosofía del *yo-sujeto*. En efecto, el actual tratamiento «estándar» del problema de la mente (o, como se formula otras veces, de la relación mente-cuerpo) se diferencia de los anteriores por partir del supuesto, al menos *metodológico* (pero también, casi siempre, *ontológico*) de que la mente es una realidad natural cuyo estudio es inseparable de la investigación sobre el mundo material. Tal como lo formula Galen Strawson:

Hay una cierta versión del monismo que es verdadera y [...], a pesar de toda la variedad del mundo, en algún sentido fundamental hay sólo una clase de estofa. [...] los sucesos mentales humanos son realizados íntegramente por sucesos físicos o son realizaciones de éstos, en donde «realizar» tiene un sentido que tendrá que analizarse. [...] la teoría de la evolución da una explicación verdadera (aunque hasta ahora incompleta) de la manera en que seres como nosotros advinieron a la existencia en un mundo físico existente con anterioridad y [...] debemos adoptar un enfoque naturalista de la cuestión de la naturaleza de la mente. En realidad, el término «naturalismo» no es más preciso que los términos «físico» o «material». Lo único que implica realmente es el rechazo de cualquier cosa que se clasifique como sobrenatural en relación con una concepción dada de lo natural. Pero no conocemos los límites de lo natural.<sup>20</sup>

La última proposición del pasaje anterior refleja la postura «agnóstica» (agnósticamente materialista) del autor, postura no compartida por la mayoría de quienes hoy estudian la mente desde un punto de vista naturalista. En general, un observador externo de la evolución contemporánea de los trabajos teóricos en este campo sacaría la impresión de que «naturalismo» es sinónimo de «físicismo reduccionista». No simplemente «reduccionismo», como suele decirse, porque éste puede practicarse en cualquiera de estos dos sentidos: reduciendo lo físico a lo mental —como, en líneas generales, hacen

---

20. Strawson, G., *Mental Reality*, Cambridge (MS), The MIT Press, 1994 (versión castellana: *La realidad mental*, Barcelona, Prensa Ibérica, 1997; pp. 19-20).

el idealismo y el empirismo radical– o reduciendo lo mental a lo físico, que es la tendencia predominante hoy en día.

Al decir reducción nos referimos a una aparente equivalencia que, en realidad, opera en el sentido siguiente: se afirma la «fusión» de ambas entidades equivalentes en una sola, a la vez que se niega toda realidad a las propiedades de una de ellas en beneficio de las de la otra. En el caso que nos ocupa, lo que se intenta habitualmente es mostrar que las propiedades específicamente mentales carecen de realidad (serían, según eso, «ilusiones») y que las únicas reales son las propiedades *no mentales*. Paradójicamente, no son los autores que más abiertamente propugnan la llamada «teoría de la identidad» (en general, los de la llamada «escuela australiana») quienes incurrir en este reduccionismo de propiedades. Así, por ejemplo, J.J.C. Smart, en «Sensations and Brain Processes», dice simplemente:

Si el dolor es idéntico a un estado *B* del cerebro, debe haber una concomitancia entre los casos de dolor y los casos de estado cerebral *B*. [...] Así, una condición necesaria de la identidad entre el dolor y el estado cerebral *B* es que las dos expresiones, 'tener dolor' y 'hallarse en el estado cerebral *B*', tengan la misma extensión.<sup>21</sup>

Tener la misma extensión no equivale a tener la misma comprensión. No sólo eso: la identidad extensional no puede deducirse analíticamente de la comprensión de ambos términos, sino que son precisos criterios empíricos que acrediten la identidad:

Hemos sostenido que cualquier teoría satisfactoria de la mente ha de admitir la posibilidad lógica de mentes sin cuerpo. [...] Si hay algo seguro en filosofía, podemos tener por seguro que [la proposición] «la mente es el cerebro» no es una verdad lógicamente necesaria. Cuando Aristóteles decía que el cerebro no era más que un órgano para refrigerar el cuerpo no se le podía ciertamente acusar de negar una verdad necesaria. Su error era de índole empírica. Así, si es verdad que la mente es el cerebro, habrá que encontrar un modelo entre los enunciados de identidad contingentes. [...] Pero si «la mente es el cerebro» es un enunciado contingente, se sigue de ahí que ha de ser posible dar explicaciones lógicamente independientes (o, alternativamente, «definiciones ostensivas») del significado de las palabras 'mente' y 'cerebro'.<sup>22</sup>

---

21. «Sensations and Brain Processes», *Philosophical Review*, 68 (1959), 141-156.

22. Armstrong, D.M., *A Materialist Theory of the Mind*, Londres, Routledge, 1963 (1993), pp. 76-77.

De hecho es aquí donde se apoyan los dualistas residuales para defender su posición. Basándose en el «reverso» negativo del principio de identidad de los indiscernibles, formulado por Leibniz,<sup>23</sup> algunos autores, empezando por Descartes en el *Discurso del método*, construyen el siguiente argumento:

- (D1)
1. *Es posible que yo no tenga cuerpo* (duda metódica)
  2. *No es posible que yo no tenga mente* (evidencia del cogito)
  3. *Luego mi cuerpo y mi mente son distintos*

El propio Leibniz rechaza esa utilización de su principio argumentando que en la primera premisa hay un paso implícito (e ilegítimo) de la posibilidad epistémica a la ontológica, o de la duda a la negación («si dudo de mi cuerpo es que no tengo cuerpo»). No obstante, es posible reformular el argumento de manera que se evite la equivocidad denunciada por Leibniz, a saber:

- (D2)
1. *Es posible que yo sea una mente pero no tenga cuerpo*  
(duda metódica)
  2. *Si la mente es idéntica al cuerpo, necesariamente todo x que sea mente será también cuerpo* (aplicación de la teoría de la identidad de Saul Kripke)<sup>24</sup>
  3. *Luego el cuerpo y la mente no son idénticos*<sup>25</sup>

Cabe asimismo la posibilidad, siguiendo a Descartes,<sup>26</sup> de formular esta otra aplicación del principio de identidad de Leibniz que tiene además la estructura de un silogismo aristotélico del modo *Cesare* de la segunda figura:

- (D3)
1. *Ninguna mente es divisible*
  2. *Todo cuerpo es divisible*
  3. *Luego el cuerpo y la mente no son idénticos (ningún cuerpo es una mente)*

El problema con todos estos argumentos es, tal como advierte Armstrong en la obra arriba citada, el siguiente:

---

23. El principio de identidad de los indiscernibles establece que: «si dos cosas son idénticas, han de tener en común todas sus propiedades»; si negamos el consecuente, habremos de hacer otro tanto con el antecedente: dos cosas que tengan siquiera una propiedad no común han de ser diferentes.

24. Kripke, S., *Naming and Necessity*, Cambridge (MS), Harvard University Press, 1972 (1993), pp. 97-105.

25. Bealer, G., «Mental properties», *Journal of Philosophy* 91 (1994), pp. 185-208. El propio Bealer, sin embargo, considera el argumento no concluyente.

26. *Meditaciones metafísicas*, meditación sexta [33].

Si tiene sentido decir «La mente es el cerebro», ha de ser posible tratar los términos 'mente' y 'cerebro' de la misma manera [que el enunciado «El gen es una molécula de AND»<sup>27</sup>]. La palabra 'cerebro' no plantea ningún problema. Está claro que es posible explicar su significado de manera casi ostensiva. El problema lo plantea la palabra 'mente'. ¿Qué explicación verbal o «definición ostensiva» podemos dar del significado de esta palabra sin presuponer una toma de postura que nos aleje de la concepción fisicista del mundo?<sup>28</sup>

En efecto, aquí está el quid de la cuestión. Armstrong lo apunta, pero acaba ignorando la fuerza de su propio argumento, al creer que es posible la redefinición de 'mente' como el *nexo causal entre los estímulos y las respuestas del organismo humano*.<sup>29</sup> Redefinición que le permite, sin dificultad alguna, identificar ese nexo causal, y por tanto la mente, con el sistema nervioso central. Aquí es, en definitiva, donde volveremos a la hora de formular nuestra propia tesis al respecto.

Las posiciones de los partidarios de la teoría de la identidad se endurecen hasta el extremo en el caso de los llamados «elimitativistas» (o, mejor, «eliminacionistas»). Esta corriente, representada paradigmáticamente por los esposos canadienses Patricia y Paul Churchland,<sup>30</sup> sostiene que, a través del estudio del cerebro y sus ramificaciones nerviosas con arreglo a modelos computacionales de tipo especial —no secuenciales, como en los ordenadores convencionales, sino concebidos para trabajar en paralelo como las neuronas («procesamiento distribuido en paralelo», PDP)— es perfectamente factible obtener una nueva visión de los procesos mentales que permita «eliminar» el actual vocabulario mentalista, carente de rigor, y sustituirlo por otro, rigurosamente científico, en que expresiones del tipo 'creo que', 'deseo que', correspondientes a las llamadas «actitudes proposicionales», cederán su lugar a descripciones precisas de la actividad de determinados grupos neuronales. Consideran esos autores que, de la misma manera que la teoría geocéntrica fue abandonada al descubrirse que la tierra y los demás planetas orbitan en torno al sol, así también lo que ellos llaman «psicología popular»<sup>31</sup> y que conciben como una teoría de

27. Es decir, se ha de poder explicar el significado de cada uno de los términos de manera lógicamente independiente como paso previo a determinar si se refieren a la misma cosa.

28. Armstrong, *Op. cit.*, p. 77.

29. *Ibíd.*, p. 78.

30. Véase, del segundo, *Scientific Realism and the Plasticity of Mind*, Cambridge (MS), Cambridge University Press, 1979, y *The Engine of Reason, the Seat of the Soul*, Cambridge (MS), The MIT Press, 1995; y de la primera, *Neurophilosophy: Toward a Unified Science of the Mind-Brain*, Cambridge (MS), The MIT Press, 1986, y *The Computational Brain* (con T.J. Sejnowski), Cambridge (MS), The MIT Press, 1992.

31. Que es, por cierto, el título de la obra magna de Wilhelm Wundt (1910), máximo exponente, junto con William James, de la psicología introspectiva de finales del XIX y principios del XX.

la mente implícita en nuestro uso lingüístico convencional, quedará relegada al más absoluto olvido en cuanto tengamos listos los nuevos modelos descriptivos neuro-computacionales.

Se trata, obviamente, de un proyecto radical que supone una enorme simplificación del problema: no hay ya que esforzarse en buscar correlaciones entre nuestras actitudes proposicionales y nuestras pautas de actividad neuronal; basta con mostrar cuáles son esas pautas y describirlas con lenguaje puramente extensional de tercera persona, eliminando sin más la «teoría» implícita en la expresión de actitudes proposicionales. Pero la simplificación cae aquí en el simplismo: el eliminacionismo topa con dificultades tan obvias como la de conciliar la determinación última del significado en función del contexto (en sentido amplio, no meramente lingüístico) con la unicidad y regularidad de los procesos de excitación neuronal considerados en sí mismos, despojados de toda descripción «mentalista».<sup>32</sup> Por no hablar del problema que plantea eliminar de la jerga científica y filosófica términos como 'verdadero', 'falso', etc.

Pero el enfoque naturalista de la mente que seguramente cuenta con más adeptos es el llamado «funcionalismo», cuya tesis principal, enunciada en su momento por Jerry Fodor,<sup>33</sup> es la siguiente: la distancia aparentemente insalvable que separa el lenguaje ordinario acerca de los procesos mentales y su implícita teoría de la mente, por un lado, y los complejos y aún mal conocidos procesos electroquímicos que tienen lugar en las fibras nerviosas del cerebro y el sistema central, por otro lado, se reduce considerablemente si postulamos la existencia de un eslabón intermedio, una especie de «lenguaje de la mente»<sup>34</sup> capaz de traducir los distintos procesos físicos que tienen lugar en el cerebro humano a unos esquemas que serían las matrices de todas nuestras representaciones y experiencias, incluida la comprensión y el uso del lenguaje. En realidad, ese eslabón intermedio desempeñaría la función que los programas o aplicaciones informáticas desempeñan en los ordenadores. Y de la misma manera que el usuario de un ordenador no «ve» el funcionamiento interno de

---

32. Objeciones de este tipo las plantea, por ejemplo, Davidson en relación incluso con versiones reduccionistas más moderadas, como la de Smart, con su pretensión de que sea posible descubrir leyes psicofísicas estrictas que establezcan correspondencias biyectivas entre procesos neuronales y significados.

33. Véase su *The Language of Thought*, Scranton (PA), Thomas Y. Crowell, 1975 (versión castellana: *El lenguaje del pensamiento*, Madrid, Alianza Editorial, 1984); así mismo: *Psycho-semantic: The Problem of Meaning in the Philosophy of Mind*, Cambridge (MS), The MIT Press, 1987.

34. En la jerga especializada se lo designa, a imitación de las denominaciones de los idiomas conocidos, con el nombre de *mentalés* (lenguaje de la mente). La hipótesis del mentalés tiene uno de sus puntos de apoyo en las teorías lingüísticas de Noam Chomsky, con su postulado (anticipado ya por los antiguos estoicos) de la existencia de una facultad innata y específica del lenguaje.

su máquina, los flujos electromagnéticos que circulan en el interior de innumerables y enrevesados microcircuitos, sino sólo la información gráfica, inteligible por él o ella, que aparece en la pantalla, así también ocurre de hecho con nuestra actividad consciente: no tenemos la más remota noción de la estructura y actividad electroquímica de nuestras neuronas, ni falta que nos hace, para sentir emociones, entender y articular oraciones; esos «estados mentales» consisten en unas determinadas pautas o reglas de funcionamiento de naturaleza relacional algorítmica, que manifiestan por sí mismas cómo sentimos, pensamos, deseamos y nos expresamos; pautas que podrían, además, darse de la manera que se dan con total independencia de cuál sea el material de que están hechos los mecanismos materiales que utilizan, de la misma manera que las reglas del ajedrez no dependen de si las piezas son de un tamaño absoluto u otro (sí cuenta, obviamente, el tamaño relativo que permita distinguirlas) o de si están hechas, por ejemplo, de madera, metal o plástico. Lo cual abre, obviamente, un amplio campo a la investigación en la llamada «inteligencia artificial».<sup>35</sup> En efecto, el funcionalismo, al menos en la variante defendida en sus primeros tiempos por Putnam,<sup>36</sup> hace suya una analogía según la cual la mente es al cerebro lo que el programa informático (el *software*) es a la máquina que lo ejecuta (el *hardware*). Las ventajas de este planteamiento con respecto al eliminacionismo son evidentes:

a) Legítima (al menos en la versión del funcionalismo propugnada por Fodor) nuestro vocabulario mentalista tradicional, para el que nadie ha encontrado todavía sustitutos aceptables (sustitutos, por ejemplo, que superen ese vocabulario en flexibilidad y economía de recursos).

b) No elimina datos del problema para facilitarse la solución, como hace el eliminacionismo y, en parte, también el conductismo.

c) Admite y subraya, frente al epifenomenismo,<sup>37</sup> el papel causal de los estados mentales, sin echarse en brazos del dualismo (aunque algún funcionalista atípico, como David J. Chalmers, sostiene posturas que rozan el dualismo).

---

35. La investigación en este campo fue iniciada por quien hasta hace poco, al menos, era su figura más señera y su apóstol más entusiasta, Marvin Minsky, quiene en colaboración con John McCarthy puso en marcha en 1959 el Proyecto de inteligencia artificial del M.I.T. Un año más tarde, Hilary Putnam publicaba su célebre artículo «Minds and Machines» (*Mentes y máquinas*), verdadero manifiesto inaugural del funcionalismo «computacionista».

36. Quien luego, sin embargo, cantaría la palinodia retractándose de sus tesis funcionalistas en obras como *Representation and Reality* (Cambridge (MS), The MIT Press, 1988).

37. Son epifenomenistas aquellas teorías de lo mental que niegan a los estados mentales, en cuanto tales, cualquier poder causal sobre nuestros actos, reservando dicho poder en exclusiva para los procesos físico-químicos de nuestro cuerpo, respecto de los cuales lo mental es un sim-

Claro es que hay autores en la periferia del funcionalismo, como el muy prolífico Daniel Dennett,<sup>38</sup> a los que podríamos llamar «eliminacionistas con piel de funcionalistas». En efecto, proponiendo aparentemente caminos para establecer un «algoritmo» que permita pasar sin pérdida de grupos de conexiones neuronales a hechos de conciencia, acaban negando la existencia de ésta como tal, con sus propiedades «experienciales», su peculiar contenido cualitativo (los *qualia*, como se dice en la jerga especializada), que cualquiera ve, intuitivamente, que no puede sustituirse por descripciones externas al sujeto consciente, como a ningún ciego puede sustituirse la experiencia visual de una puesta de sol mediante la lectura de descripciones literarias del ocaso. Dennett, concretamente, habla en los últimos tiempos de «heterofenomenología» como método de análisis «objetivo» de la experiencia.<sup>39</sup>

Como señala con especial agudeza John R. Searle en su crítica del funcionalismo, especialmente en su versión computacionista conocida como «Programa fuerte de inteligencia artificial» (abreviado, «IA fuerte»), el hecho de seguir fielmente las instrucciones de un programa para el manejo de los signos de un lenguaje no equivale en absoluto a *entender* dicho lenguaje. El argumento empleado a tal fin por Searle ha alcanzado gran notoriedad y se lo conoce como «la Habitación China». En síntesis: el propio Searle estaría dispuesto a someterse al experimento de responder con ideogramas chinos a preguntas formuladas por el mismo medio, a condición de disponer de unas tablas que correlacionaran todas las preguntas posibles con las posibles respuestas; salvo error accidental, Searle «funcionaría» perfectamente como interlocutor de una comunicación en chino, sólo que, según confesión propia, Searle no sabe una palabra de chino. En sus propias palabras:

La tesis de la IA fuerte no es que un computador pueda «generar» o tener estados mentales como propiedades emergentes,<sup>40</sup> sino que el *programa realizado, por sí mismo, es constitutivo de poseer una mente; que poseer*

---

ple acompañamiento o «sombra» de la verdadera actividad, que es la corpórea. Suele, por ello, recibir también el nombre de *paralelismo psicofísico*. Una versión extrema, aunque no exenta de fecundidad teórica, de esta concepción sería la célebre doctrina leibniziana de la *armonía preestablecida*.

38. Véase sobre todo: *Consciousness Explained*, Boston, Little, Brown & Co., 1991 (versión castellana: *La conciencia explicada: una teoría interdisciplinar*, Barcelona, Paidós, 1995).

39. Una brillante exposición de este método por el propio Dennett tuvo lugar en la Facultad de Filosofía de la UB en el primer semestre de 2001, organizada por el grupo de investigación en filosofía del lenguaje y ciencia cognitiva LOGOS, dirigido por el profesor Manuel García Carpintero.

40. Se suelen llamar «emergentes» aquellas propiedades adquiridas por (o manifestadas en) un objeto como consecuencia de sufrir éste una serie de cambios ninguno de los cuales, por separado, permite predecir la aparición de (alguna de) esas propiedades.

*una mente consiste en eso. El programa realizado, por sí mismo, garantiza la vida mental.* Ésa es la tesis que el argumento de la Habitación China refuta. La refutación nos recuerda que el programa está definido en términos puramente sintácticos y que la sintaxis, por sí misma, no basta para garantizar la presencia de contenido mental, semántico. [...] En cualquier caso, a estas alturas me parece que el argumento de la Habitación China concede demasiado a la IA fuerte, pues concede al menos que la teoría es falsa. Ahora pienso que es incoherente. He aquí por qué. Pregúntese cuál es la razón por la que la máquina con la que estoy escribiendo ahora hace que sus operaciones sean sintácticas o simbólicas. En lo que respecta a su física, esa máquina no es sino un circuito electrónico muy complejo. El hecho que convierte sus impulsos eléctricos en simbólicos pertenece a la misma clase de hechos que convierten en símbolos las marcas de tinta estampadas en las páginas de un libro: nosotros hemos diseñado, programado, impreso y manufacturado esos sistemas; por eso podemos tratar y usar esas cosas como símbolos. En una palabra, la sintaxis no es intrínseca a la física del sistema, sino que está en el ojo del observador. Salvo en los pocos casos de agentes conscientes que proceden a una computación (sumando 2+2 para obtener 4, por ejemplo), la computación no es un proceso intrínseco a la naturaleza, como lo son, en cambio, la digestión o la fotosíntesis, sino que existe sólo en relación con algún agente que da una interpretación computacional a la física. De modo que la computación no es intrínseca a la naturaleza, sino que es relativa al observador o al usuario.<sup>41</sup>

## El conocimiento de la mente presupone la mente

La segunda parte del pasaje de Searle arriba citado da la clave para entender dónde reside el verdadero problema de la filosofía de la mente: en su insoslayable carácter *autorreferente*. En efecto, el conocimiento de la mente (genitivo objetivo) sólo puede ser conocimiento de la mente (genitivo subjetivo). Dicho sin retruécano: sólo la mente puede conocer la mente. Pero, a la vez, no puede conocerse como tal. Es esa circularidad de la que no podemos salir la que convierte en misterio un fenómeno natural como sin duda es la actividad cognoscitiva. De hecho, es ése el *único misterio* con el que topa la filosofía (la ciencia, propiamente, no topa con él: pasa a su lado, porque de entrada y por principio no busca vivencias sino evidencias: éstas, aunque sólo se presentan a la conciencia, pueden observarse haciendo abstracción del observador; las vivencias, contenido directo de la conciencia, sólo pueden vivirse, y la filosofía debe ocuparse también de ellas, porque la filosofía no sería fiel a sí misma si dejara de entrada algo fuera de la órbita de sus intereses).

---

41. John R. Searle, *The Mystery of Consciousness*, Nueva York, The New Review of Books, 1997, pp. 26-27 de la versión castellana: *El misterio de la conciencia*, Barcelona, Paidós, 2000.

Digo que la naturaleza y función de la mente es el único *misterio natural* porque es el único fenómeno de la naturaleza que no podemos observar con *naturalidad*, es decir, con perspectiva: no hay perspectiva posible para observar —desde el único observatorio posible, que es la mente— la mente en ejercicio; al menos si lo que se quiere observar realmente es su ejercicio y no simplemente las condiciones materiales de éste. Y aquí la diferencia entre observación «desde fuera» y observación «desde dentro» (introspección) es finalmente irrelevante. La observación externa sólo puede descubrir, en el mejor de los casos, el funcionamiento sensorialmente perceptible de la microestructura material en que se realiza la actividad mental; en el peor de los casos, se ha de limitar a registrar el comportamiento «macroscópico» del cuerpo. Pero la introspección, aunque capta y registra sin duda todo el cúmulo de experiencias vividas que constituyen el flujo de cada conciencia individual, no es menos ignorante acerca de la estructura profunda, la «tramoya» de ese espectáculo que desfila ante ella sin mostrar nada más que una superficie fenoménica tras la que suponemos la presencia de objetos independientes de nuestra observación a la vez que nos reconocemos incapaces de concebirllos si no es como potencialmente observables. Y es que, en el fondo, como señala Roger Penrose en sus escritos sobre la mente,<sup>42</sup> nuestra ignorancia sobre la naturaleza de la mente es indisociable de nuestra ignorancia sobre la naturaleza profunda de la materia, naturaleza que no se deja apresar sin más en las fórmulas matemáticas de la física newtoniana o einsteiniana.

Porque eso es así, es decir, porque la introspección nos deja en el fondo tan insatisfechos como la observación externa, la filosofía contemporánea de la mente, bajo la guía de Wittgenstein y de los psicólogos conductistas, ha sido predominantemente «externalista», ya que en ese enfoque la filosofía parece contar al menos con la compañía de la ciencia y compartir con ella sus gozos y, sobre todo, sus penas. En último término, podría pensarse en un reparto del trabajo entre ciencia y filosofía que asignara a aquélla la tarea de observar «desde fuera» la plasmación neurofisiológica de la actividad mental y a ésta una tarea introspectiva de registro y análisis de experiencias mentales.

Pero eso sería renunciar lisa y llanamente a hacer filosofía de la mente y sustituir ésta por un híbrido de neurociencia más o menos rigurosa y psicología descriptiva totalmente carente de rigor. Mientras el científico se dedicara a buscar la «fórmula de la sopa» renunciando de entrada —según el dicho de Einstein— a que dicha fórmula supiera ella misma a sopa, el filó-

---

42. Penrose, R., *The Emperor's New Mind*, Oxford, Oxford University Press, 1989 (versión cast.: *La nueva mente del emperador*, Madrid, Mondadori, 1991); *Shadows of the Mind*, Oxford, Oxford University Press, 1994.

sofo se entregaría a saborear distintos tipos de sopa para tratar de descubrir qué relaciones guardan entre ellos y si participan o no de una misma «sopa universal».

Pero tanto la fórmula de la sopa como su sabor son objetos de la mente, no la mente misma. Ahora bien, ¿acaso es posible investigar la mente misma? Los antiguos neoplatónicos y otros muchos filósofos posteriores, como Descartes, pensaron que esa investigación podía llevarse a cabo sin esfuerzo porque ella, el *noûs*, era transparente a su propia mirada. Muchos contemporáneos se dejan engañar por una imagen especular de ese mismo espejismo: la mente, como para los neoplatónicos, se identifica con su objeto, al menos cuando ese objeto es el sistema nervioso central o el «programa» con arreglo el cual funciona. Para aquéllos, la mente absorbe el objeto; para éstos, el objeto absorbe la mente. El resultado es en ambos casos el mismo: una pretendida transparencia de la mente a sí misma que en realidad es la autodisolución de lo mental.

Porque la mente no es nada más (ni nada menos) que la referencia de ese insignificante pronombre de primera persona desde cuyo punto de vista sentimos, deseamos, pensamos y hablamos. Referencia, por cierto, que no tiene ubicación en ninguno de los fenómenos de la experiencia, ni externa ni interna, por lo cual no merece siquiera el nombre de referencia. Referencia que, aunque la haya, cambia constantemente de asiento en el curso de un diálogo. Referencia que remite siempre, en último término, a algo que está más allá de nuestro estado psicológico de cada momento, a un trasfondo que ni siquiera está todo él dado, realizado en el pretérito y el presente, sino que incluye la posibilidad y el futuro. Referencia, en fin, que parece circunscribirse al espacio del cuerpo de quien habla pero que en realidad se extiende a todos los espacios que el hablante ocupa, ocupó y ocupará y que, por tanto, habita con más propiedad en el tiempo, un tiempo sin límites definidos.

Eso y no otra cosa es la mente, al menos la mente humana.<sup>43</sup> Por supuesto, ese centro de perspectiva está *constituido*<sup>44</sup> por una base material (un cuerpo y, más precisamente, dentro de él, un cerebro con todas sus ramificaciones nerviosas). Pero la relación de constitución es asimétrica y no

---

43. En los animales, al menos en los mamíferos, es más que probable que no falte tampoco un «centro de perspectiva» en torno al cual se organice la experiencia. Pero en el ser humano ese centro puede a su vez simbolizarse y convertirse en objeto de atención por sí mismo, hecha abstracción de la experiencia: es este *yo* totalmente explícito el que Lynne R. Baker considera propiamente centro de «perspectiva de primera persona» (véase su *Persons and Bodies*, Cambridge, Cambridge University Press, 2000).

44. En el elaborado sentido de 'constitución' expuesto en el libro de L.R. Baker mencionado en la nota anterior, acepción que incluye no sólo las propiedades intrínsecas del «material» constituyente, sino también sus propiedades relacionales.

debe confundirse con la relación de identidad. Así, pues, la perspectiva de primera persona está constituida en cada caso por un cuerpo pero *no se identifica* con él.

Esto nos lleva de la mano al pasaje de la obra de David Armstrong citado más arriba, donde reconoce que no parece posible explicar el término 'mente' sin recurrir a expresiones específicamente «mentales»: «¿Qué explicación verbal o "definición ostensiva" podemos dar del significado de esta palabra sin presuponer una toma de postura que nos aleje de la concepción fisicista del mundo?» En efecto, más allá de afirmar que está constituida por tal o cual sustrato material, con el cual no se identifica sin más, nada significativo podemos añadir. Salvo, quizá, una cosa: que el centro reflejo de perspectiva que abreviadamente llamamos *yo* es esencialmente un *nudo de relaciones*. Y la relación, como dice Aristóteles, es la categoría más alejada de la substancia. Una concepción, pues, de la mente como entramado de relaciones no precisa añadir nada substancial, ni siquiera cualitativo, al material constituyente (de naturaleza neuronal o de cualquiera otra, con tal de que sea capaz de soportar aquellas determinadas relaciones).<sup>45</sup> Desde este punto de vista, la hipótesis funcionalista de un nivel intermedio entre soporte físico y estados mentales resulta superflua. Pero también resulta desencaminada, en parte al menos, la crítica de Searle a la idea funcionalista de que «la mente es el programa» que ejecuta el cerebro. Esto último es inadmisiblesi se entiende por programa lo que suele entenderse en informática: un algoritmo que permite ejecutar ciertas operaciones en un determinado orden con vistas a obtener unos resultados computables (es decir, previsibles). Pero no lo es tanto si se entiende como un conjunto *abierto* de operaciones de diversa índole que incluyen particularmente eso que vivimos como *experiencia*. No hace falta ningún nivel intermedio de tipo algorítmico para explicar los procesos mentales como constituidos por procesos neuronales. El problema que le planteaba a Armstrong la asimetría en la definición, respectivamente, de 'cerebro' y 'mente' no es tal problema, es precisamente la vía de solución: la mente es un determinado cúmulo de relaciones que el cerebro puede establecer dentro de sí, con el resto del cuerpo y, a través de éste, con el mundo exterior. Es lógico que haya asimetría entre el concepto de un órgano y el de su función.

---

45. No hay motivo para rechazar la posibilidad de construir artificialmente verdaderas inteligencias: si la evolución «ciega» lo ha hecho, ¿por qué no habría de poder hacerlo una estrategia consciente? Lo que ocurre es que, hoy por hoy, lo que se sabe sobre el funcionamiento de las neuronas y sus asociaciones es una ínfima parte de lo que deberíamos saber para reproducir, incluso con otros materiales, todas las relaciones constitutivas de un hecho experiencial. Además, con toda seguridad, esas relaciones formarán un conjunto abierto, no computable, sin lo cual parece inconcebible el aprendizaje.

¿Por qué insisto en el aspecto relacional? Porque sólo una relación se puede «superponer» a cualquier tipo de realidad material sin añadirle materialidad, requisito imprescindible para no recaer en el dualismo, del tipo que sea (muchos autores contemporáneos, en efecto, huyendo del dualismo de substancias, caen en el dualismo de atributos o propiedades, o hablan de «propiedades emergentes», lo que, en coincidencia con Colin McGinn<sup>46</sup> y Thomas Nagel,<sup>47</sup> considero una forma subrepticia de dualismo ontológico «por etapas» o, peor incluso, de «creacionismo»). De ahí el fondo de verdad que creo hay que reconocerle al conductismo, sobre todo en su vertiente crítica de las concepciones substancialistas de la mente.

Podría parecer que las relaciones tienen tan poca entidad intrínseca (en principio, ninguna) que reducir la mente a un conjunto de ellas es tanto como recaer en el epifenomenismo. A esto hay que responder que la asimetría substancia/cualidad – relación ha sido exagerada por la tradición esencialista-fixista que arranca de Platón (o de una cierta interpretación dominante de su filosofía). En realidad creo que la «substancia de la substancia» ha de considerarse ella misma relacional: lo es, si nos atenemos a la concepción hilemórfica aristotélica de las substancias materiales, en lo que respecta a la materia, que es pura *relacionalidad espacio-temporal*, y de manera algo diferente en lo que respecta a la forma, que es *estructura ordenada* y, por tanto, expresable como un conjunto de relaciones de orden entre los elementos estructurales. Según esto, la matriz del sistema categorial de Aristóteles sería en realidad la relación, pues todas las demás categorías se pueden definir a partir de ella.

Vista así, la asimetría mente-cuerpo resulta mucho menos escandalosa, y se hace mucho más fácil encontrarle acomodo a la mente en una metafísica monista no idealista.

De todos modos, siendo la mente el *yo*, sujeto último de toda representación, centro de perspectiva del que no podemos abdicar por más que recurramos a subterfugios infantiles como intentar traducir todo nuestro lenguaje «mental» a expresiones en tercera persona,<sup>48</sup> no hay miedo de que lo mental aparezca en competencia con lo material a la hora de explicar los hechos del

46. McGinn, C., 'Can We Solve the Mind-Body Problem?', *Mind* 98 (1989).

47. Nagel, T., 'Panpsychism', en: *Mortal Questions*, Cambridge, Cambridge University Press, 1979.

48. Es «infantil» en sentido literal, pues son los niños los únicos humanos que, antes de aprender el uso de los pronombres personales, se refieren espontáneamente (no «aviesamente», como ciertos filósofos) a sí mismos mediante un sujeto nominal y un verbo en tercera persona. Pero, por supuesto, si «se ven a sí mismos» en tercera persona es porque, aun antes de dominar la expresión verbal correspondiente, están ya *situados* en la primera.

mundo en los que intervenimos los humanos. Porque, como dice Wittgenstein en la proposición 5.641 del *Tractatus*:

Existe, pues, realmente un sentido en el que en filosofía puede tratarse no-psicológicamente del yo. El yo entra en la filosofía por el hecho de que «el mundo es mi mundo». El yo filosófico no es el hombre, ni el cuerpo humano, ni el alma humana, de la que trata la psicología, sino el sujeto metafísico, el límite; no una parte del mundo.

## NORMATIVA PARA LA COLABORACIÓN EN «CONVIVIUM»

### A) *Normas generales sobre los trabajos, su recepción y publicación:*

1. CONVIVIUM publicará tres tipos de trabajos: «estudios», «notas o discusiones» y «reseñas». Éstos podrán estar escritos en cualquiera de las lenguas latinas, en inglés o en alemán.
2. Los autores de las colaboraciones deberán enviar a la redacción de la Revista -o a cualquiera de los miembros de su Comité de Redacción- dos copias de su trabajo escrito pulcramente a máquina en hojas de tamaño DIN-A 4 por una sola cara, con buen margen, con interlineación a doble espacio, y de una extensión que no sobrepase, en general, las 35 páginas (o sea, alrededor de las 14.000 palabras si la redacción es en castellano); además, grabado en un disquete; sistemas: Word 95 o posteriores.
3. Cuando el trabajo sea del tipo «estudio», el autor incluirá un resumen del mismo que no exceda de las 150 palabras y que se publicará precediendo al cuerpo del artículo.
4. Junto con las 2 copias del original de su trabajo, los autores enviarán a la Redacción los datos relativos a sus titulaciones académicas, cargos y docencia -si los hubiere-, dirección actual y n.º de teléfono.
5. En cuanto obre en su poder un trabajo, la Redacción notificará a su autor la recepción del mismo.
6. Los originales recibidos no serán devueltos, pero la Redacción se reserva el derecho de aceptarlos o no en orden a su efectiva publicación según su conveniencia y oportunidad para cada número de la Revista. Con este fin, al recibir cada trabajo, la Redacción encargará a dos lectores o revisores cualificados e independientes que enjuicien los méritos del mismo.
7. En el caso que, cumplidos los anteriores requisitos, un trabajo vaya a formar parte de uno de los números de la Revista, la Redacción notificará a su autor la fecha previsible de su publicación.
8. La Redacción no se solidarizará en ningún caso con las opiniones expuestas en los trabajos que en la Revista se publiquen, y sobre este particular no mantendrá correspondencia de ningún género.
9. Los autores recibirán gratuitamente 20 separatas de los trabajos del tipo estudio, 10 de las notas o discusiones y 5 de las reseñas.

### B) *Normas específicas sobre presentación de originales*

10. Para una mayor claridad expositiva, se aconseja el uso de suficientes divisiones y apartados en el texto.

11. Para las citas muy largas que se incluyan en el texto se aconseja emplear párrafos en letra pequeña, particularidad que se indicará para la imprenta poniendo una línea vertical, a lo largo de toda la extensión de la cita, en el margen izquierdo.
12. Toda abreviatura, así como toda referencia textual o bibliográfica, deberá figurar como nota al pie de página.
13. En el caso de texto o términos griegos, estos habrán de ir transliterados según los convencionales al uso.
14. Las llamadas a las notas a pie de página se escribirán en el texto del cuerpo del trabajo con sucesivos números volados y sin paréntesis. El contenido de las notas a que tales números remitirán se dará a la Redacción en páginas especiales que vayan al final de cada trabajo, procurando que su numeración corresponda exactamente con la de las respectivas llamadas.
15. En las notas a pie de página, las citas deberán ser completas y exactas; se las redactará del siguiente modo:
  - Para libros: Apellido(s) del autor, iniciales de su nombre, título del libro *subrayado*, lugar de la edición, editorial, año de la edición, página o páginas citadas.
  - Para citar artículos de revista: Apellido(s) del autor, iniciales de su nombre, título entre comillas del artículo, nombre de la revista *subrayado*, número del volumen, año entre paréntesis, paginación del artículo o número(s) de la(s) concretamente citada(s).
16. Se usarán las comillas para citas de textos -cualquiera que sea el idioma en que se hagan- y para los términos empleados en sentido poco frecuente o con intención especial. La letra cursiva, que se indicará para la imprenta mediante subrayado, se reservará para destacar dentro del texto determinadas palabras o frases y también vocablos extranjeros.
17. Las listas bibliográficas que se juzgue preciso o conveniente adjuntar irán ordenadas alfabéticamente por autores y, si es posible, repartidas en secciones según los temas o materias.
18. Las instrucciones especiales para el impresor deberán encerrarse en círculos puestos al margen, a ser posible con una grafía que se destaque por su color.
19. Los originales que no se ajusten del todo a estas normas, supuesto que la Redacción los estime publicables, quedarán expuestos a graves retrasos de publicación; pero, si se le pide al autor que los enmiende conforme a estas normas, no será sobre los ya enviados por él, sino que la Redacción habrá de recibir nuevo ejemplar doble del original así enmendado, La Redacción sólo mantendrá correo de ida y vuelta para las pruebas de imprenta, no para que se corrijan originales defectuosos según los presentes requisitos.
20. Cfr. *supra*: 2) acerca del envío en soporte informática: Word 95 o ss.