

La cuestión del “lugar cerebral” del espíritu: fisicalismo, reduccionista e internalismo en *El hombre neuronal**

Pascale Gillot**

DOI: 10.22395/csye.v5n10a12

En *El hombre neuronal*, Jean-Pierre Changeux desarrolla, a partir de su práctica teórica de neurobiólogo, un programa tocante a los dominios tradicionales de la psicología y de la filosofía del espíritu. Este programa, que se designa con la denominación reduccionismo fisicalista, apunta generalmente a la explicación del conjunto de los fenómenos –en particular de los fenómenos mentales– en el vocabulario de las ciencias de la naturaleza, a la reducción explicativa o epistemológica de tales fenómenos. De esta manera puede entenderse, a título ejemplar, el proyecto de establecimiento de una “pasarela” entre lo mental y lo neuronal¹. Se trata, en otros términos, de constatar la existencia necesaria de una “fosa explicativa” que separa irremediamente lo psicológico de lo neurobiológico.

Tenemos aquí en este caso una apuesta bastante parecida a la que hicieron, veinte años después, filósofos del espíritu tan diferentes como Jaegwon Kim o Daniel Dennett, con el programa de una renovación de la teoría del acaecimiento y con el de una ciencia de la conciencia, en ruptura con otra tesis que atraviesa la *philosophy of mind* hoy, la de un supuesto “misterio” esencial de la conciencia².

Sin embargo, yo querría insistir sobre la especificidad de ese reduccionismo característico de *El hombre neuronal*, que tiende a la reducción epistemológica de la psicología y de la filosofía del espíritu a la neurobiología, buscando así

* Tomado de Michel Morange, Francis Wolff y Frédéric Worms (dirs.), *El hombre neuronal, treinta años después. Diálogo con Jean-Pierre Changeux*. París: Rue d’Ulm/Presses de l’École normale supérieure, 2016. pp. 61-75. Traducción del francés al español de Luis Alfonso Palau Castaño, Medellín, septiembre de 2016.

** Profesora asociada de la Universidad François Rabelais de Tours.

¹ Ver en particular el cap. 5 consagrado a la teoría de los “objetos mentales”.

² En el número de los partidarios de un “misterio” irreductible de la conciencia, se puede evocar especialmente a Colin McGinn (cfr. su obra *the Mysterious flame: Conscious Minds in a material world*). En el polo opuesto de estos teóricos de la “fosa explicativa” entre lo mental y lo físico, Daniel Dennett propone por su parte, a partir de una perspectiva externalista-funcionalista, la elaboración de una verdadera “ciencia de la conciencia” (*Sweet dreams. Philosophical obstacles to a science of consciousness; Dulces sueños. Obstáculos filosóficos para una ciencia de la conciencia*. Buenos Aires: Katz, 2006).

sustituir sin residuo el modelo del “Espíritu” por el de “El hombre neuronal”. Esta especificidad se sostiene en la hipótesis general –de orden filosófico– de un lugar intracerebral del pensamiento. Esta hipótesis (aunque ella misma recuse la tesis unitarista) hace del cerebro o de la “máquina cerebral” el “continente” privilegiado de los estados mentales, definidos bajo la categoría general de la representación. Por consiguiente, me parece que se encuentran muchos ejes, o muchos postulados “internalistas”, subyacentes al proyecto de *El hombre neuronal*, es decir, a su proyecto fisicalista.

Precisemos que el internalismo designa esa teoría del espíritu que identifica de manera general los estados mentales con estados internos, en el espíritu o en el cerebro, en la medida en que –como es el caso en *El hombre neuronal*– se adhiere a la teoría de la identidad estricta entre lo mental y lo cerebral.

Examinaré, pues, un cierto número de esos postulados internalistas que operan en la obra de Jean-Pierre Changeux, que me parecen indisolubles de la perspectiva fisicalista general. Esos postulados son: en primerísimo lugar la tesis de la inscripción cerebral de lo mental; segundo y correlativamente, el representacionalismo, o dicho de otro modo: la definición general del pensamiento bajo la categoría de la representación; y tercero, la hipótesis de un lenguaje del pensamiento, hipótesis tomada de la filosofía funcionalista-cognitivista, con la que el programa de *El hombre neuronal* mantiene una relación muy ambivalente.

Simultáneamente me preguntaría si la perspectiva reduccionista defendida en ese libro (constituida por la pareja fisicalismo-intenalismo, que tiene como principio la teoría de la identidad de lo mental y de lo cerebral) tiene claramente como único adversario el paradigma “espiritualista” y dualista identificado por Changeux con la filosofía de Bergson. Se trataría de determinar si esta perspectiva no se encuentra contestada igualmente, y quizá de forma más radical, por otra “tradición” distinta de todo espiritualismo, a saber: el externalismo, la teoría de la inscripción social, lingüística y antropológica del espíritu y de los fenómenos mentales, teoría según la cual los fenómenos mentales advienen “en el mundo” y no “en el cerebro”.

La inscripción cerebral de lo mental

La perspectiva de Jean-Pierre Changeux se inserta de manera general en la tradición del “localizacionismo”, a saber, la teoría de las localizaciones cerebrales heredada de Gall y de Broca. Ella reanuda pues con la hipótesis aferente de un “lugar del pensamiento”; se trata en este caso de un lugar intracerebral, aunque incluso ese lugar esté difractado, a contrapelo de la concepción unita-

rista que es explícitamente recusada en *El hombre neuronal*. Es lo que se deriva también, en el capítulo 5 de la obra, de la caracterización general del cerebro como “continente” de los objetos mentales, definidos bajo la categoría general de la “representación”.

Pero, más específicamente, parece que el trabajo realizado en *El hombre neuronal* procede de una reactivación de la teoría clásica de la identidad (de lo mental y de lo cerebral). Esta teoría de la identidad espíritu-cerebro tiene en efecto por programa la reducción de los procesos mentales a procesos físicos, y en este caso cerebrales; postula una identidad estricta, término a término, entre acontecimientos mentales y sucesos neurofisiológicos correspondientes; por ejemplo, el estado mental “dolor” no sería nada distinto a un cierto proceso neurofisiológico determinado, ligado a “la activación de las fibras C”. Se la ha desarrollado a fines de los años 1950 por autores tales como Ullin T. Place o Jack J. Smart³, y más particularmente movilizada por la explicación de la experiencia fenoménica y de la consciencia. Esta teoría postula, pues, la existencia de “leyes-puente” como principio de la correspondencia nomológica estricta entre lo mental y lo cerebral. Correlativamente, hace de la “máquina cerebral” el sustrato necesario del pensamiento, identificada con un fenómeno intracerebral.

Ahora bien, Jean-Pierre Changeux está de acuerdo con los principales ejes de la teoría de la identidad. Recusa simultáneamente el “espiritualismo” (encarnado, según él, por la filosofía de Bergson), el conductismo (que niega la eficacia causal de estados internos para la explicación del comportamiento) y el modelo del espíritu-computador fundado, según él, en una “analogía engañosa” entre el funcionamiento cerebral y el funcionamiento de la máquina lógica. La teoría de la identidad estricta que está aquí en juego compromete no solamente la hipótesis del lugar intracerebral del pensamiento, sino que define a este bajo la categoría general (internalista) de la representación. Por ejemplo, en el examen de la noción de “objetos mentales”, reportada a la noción crucial de imagen mental, Changeux reconduce de hecho la caracterización del cerebro humano como “continente” de la actividad mental. Precisa al comienzo del capítulo 5:

[...] ¿Es el cerebro una máquina de pensar? Bergson en *Materia y Memoria* escribía que “el sistema nervioso no tenía nada que sirviera para fabricar o incluso preparar representaciones”. La tesis desarrollada en este capítulo es exactamente la contraria de la de Bergson. El encéfalo del hombre que sabemos que contiene en la organización anatómica de su corteza, representaciones del mundo que lo rodea, es también capaz de construir las y utilizarlas en sus cálculos⁴.

³ Ullin T. Place, “Is consciousness a brain process?” Jack J. Smart, “Sensations and brain processes”.

⁴ J.-P. Changeux. *El hombre neuronal*. Madrid: Espasa-Calpe, cap. cit.

El cerebro es, pues, identificado como el “lugar”, el sustrato y el principio del pensamiento bajo sus diversas modalidades: perceptos, emociones, conceptos. Este modelo filosófico, clásico, del cerebro como sede del alma se encuentra primero sistematizado en el siglo XVII en el contexto de la emergencia del mecanicismo y de la ciencia moderna. Se encuentra de ello una versión crucial en Descartes, en particular en las *Pasiones del alma*, y una primera crítica, virulenta, en Spinoza. Ahora bien, un tal modelo comanda todavía el programa científico contemporáneo de una reducción explicativa de las funciones mentales a las funciones neurales, o dicho de otro modo: el programa de una actualización de los “fundamentos biológicos” de las facultades tradicionalmente consideradas como facultades psíquicas. Ese programa –el del reduccionismo fisicalista– tiene por horizonte la construcción de esa “pasarela” entre lo mental y lo cerebral, cuya realización conducirá, quizá, contra todas las tentaciones espiritualistas, a la eliminación misma del término “espíritu”, a su evicción del vocabulario de la ciencia.

Así, el modelo de la identidad espíritu-cerebro, con sus correlatos internalistas y representacionalistas, se presenta como un requisito indiscutible del proceder científico; se encontraría necesariamente postulado por la lógica misma del progreso y del descubrimiento científicos, en el campo de la neurobiología.

Sin embargo, desde fines de los años 1950, esta teoría de la identidad se encontró rápidamente cuestionada. Piénsese en los trabajos de Herbert Feigl que desvelan la insistencia del problema del cuerpo y del espíritu en relación con la cuestión de la causalidad mental, y sugieren la pista de un “doble conocimiento”⁵. Esta pista es retomada y desarrollada por Thomas Nagel bajo la denominación de “teoría del doble aspecto”. Articulada al postulado de una irreductibilidad lógica de las propiedades mentales –en lo que concierne especialmente “a lo que eso hace de”, o también “las sensaciones en bruto”– esta conduce al rechazo del fisicalismo clásico⁶. El fisicalismo neurologista comprometido en la teoría de la identidad espíritu-cerebro parece, por lo demás, hacer abstracción de la cuestión del cuerpo propio, es decir, del cuerpo considerado en su unidad y su individualidad específica, en su interacción con el entorno, por distinción con la sola “máquina cerebral”.

Uno puede igualmente referirse a la cuestión de “la anomalía” de lo mental ligada a su carácter “holístico”, tal como fue tematizada por Donald Davidson⁷. Esta dimensión holística parece haber hecho fracasar el proyecto inicial de establecimiento de una correspondencia estricta, término a término, entre estados

⁵ Herbert Feigl, “the ‘Mental’ and the ‘physical’ ”

⁶ Thomas Nagel, “Physicalism”, in H. Feigl, W. Sellars & K. Lehrer (eds.), *New readings in philosophical analysis*.

⁷ Donald Davidson (1980). *Essays on Actions and Events. Ensayos sobre acciones y sucesos*. México: Instituto de investigaciones filosóficas de la UNAM/Barcelona: Crítica, Grijalbo, Mondadori, 1995.

mentales y estados neurofisiológicos. El carácter problemático o inencontrable de esas famosas “leyes-puente” (*bridge-laws*) se encuentra aún subrayado por un filósofo del espíritu contemporáneo, Jaegwon Kim. Este autor –sin embargo fiel a la renovación del programa fisicalista– toma nota de un cierto número de problemas remanentes en filosofía del espíritu, como el de la causalidad mental, en este caso: la eficacia causal de lo mental en un mundo físico –en relación con el problema de la “clausura causal”– o también la tesis de la realización múltiple de lo mental⁸.

La crisis del fisicalismo y de la teoría aferente de la identidad estricta ha permitido el desarrollo, contra el materialismo neurologista, del cognitivismo, que propone comparar el espíritu con una “máquina de Turing”⁹. Esta teoría cognitivista, que es ella misma una de las fuentes del desarrollo de las neurociencias, reposa sobre una filosofía funcionalista inicialmente competidora de la teoría de la identidad. En efecto, por el modelo del espíritu-computador, y por la tesis conexas de la realización múltiple de lo mental, ella se opone a la representación “neurologista” de un “lugar” o de un “teatro” específico y privilegiado de lo mental, en este caso: un lugar intracerebral.

Conviene anotar que esta tesis de la realización múltiple es una tesis crucial del cognitivismo, y originariamente de esta filosofía funcionalista elaborada por Hilary Putnam. Contra la teoría de la identidad, que hace del pensamiento el privilegio exclusivo de una organización biológica específica, el cerebro humano, la teoría funcionalista plantea que los estados mentales son estados funcionales, definidos por una secuencialidad regulada, la relación causal o el procedimiento determinado que norma su sucesión. Estos estados mentales pueden, pues, realizarse en soportes físicos diferentes, por ejemplo, de tipo neurofisiológico, pero también mecánico o electrónico. Así la actividad mental, como la actividad de cálculo de la máquina de Turing, no depende de la naturaleza particular del sustrato material en el que ella se encuentra implementada. La tesis de la realización múltiple es invocada por el modelo computacional del espíritu, el que autoriza la comparación del espíritu con la máquina de cálculo –el computador– considerada bajo el aspecto de su actividad lógica, algorítmica, y no bajo la de su construcción o de su constitución física particular.

Si la teoría de la identidad de lo mental y de lo cerebral no constituye ya, desde entonces, el modelo exclusivo del pensamiento y de la intencionalidad, uno puede preguntarse si el desarrollo de las neurociencias desde hace treinta años, ha permitido efectivamente el establecimiento de esta pasarela entre lo

⁸ Jaegwon Kim. *Mind in a physical World. An Essay on the Mind-Body Problem and Mental Causation*. Cfr. en particular el final del cap. 1, § 5.

⁹ Cfr. Hilary Putnam, “Minds and machines”, in *Mind, Language and Reality*, pp. 362-385.

mental y lo neural invocada en *El hombre neuronal*. En efecto, se puede suponer que la realización del programa fisicalista e internalista, que busca sustituir el antiguo modelo del “Espíritu” (y más ampliamente una tradición espiritualista cuyo paradigma sería la filosofía de Bergson) por el modelo neuronal, tropieza con la crisis persistente del fisicalismo en filosofía del espíritu, a pesar de las tentativas recientes de reactivación de un tal fisicalismo¹⁰.

Parece, además, que la teoría fisicalista de la identidad tampoco da razón de ser de la gramática específica de lo mental, ligada a la especificidad de la experiencia en primera persona. Sin mencionar aquí el célebre problema de los qualia, se puede simplemente hacer referencia a lo que Wittgenstein identifica como la asimetría característica de los verbos psicológicos en primera persona, ligada a la infalibilidad específica de tales enunciados en primera persona, que los distingue conceptual, o gramaticalmente, de los enunciados en tercera persona. Por ejemplo, se puede evocar la distinción gramatical entre el enunciado “me duele” y el enunciado “le duele”¹¹.

El representacionalismo y la concepción de la interfaz: “lo esencial ocurre dentro de la máquina”

El representacionalismo consiste en la definición general de los procesos mentales (perceptos, afectos, conceptos) bajo la categoría general de la representación y hacer un uso determinante de la noción de “imagen mental”. Se inscribe en una tradición empirista de la que se reclama, por lo demás, explícitamente Jean-Pierre Changeux. Para esta tradición, el objeto mental es un duplicado debilitado de la sensación, y los conceptos son el resultado de un proceso de abstracción de los datos sensoriales.

El representacionalismo está, pues, así operando en la “concepción de la interfaz”, según la cual la relación del espíritu con el mundo es indirecta y supone la mediación de impresiones o de imágenes mentales que residen en el espíritu, en este caso en el cerebro.

La noción de imagen mental ocupa un lugar crucial en el dispositivo argumentativo de *El hombre neuronal*, y especialmente en el capítulo 5 que propone una teoría de la formación de los “objetos mentales” en general. La existencia de las “imágenes mentales” se encuentra en efecto postulada, a título de evidencia, por el programa de *El hombre neuronal*, con referencia al desarrollo de la “neuroimagenografía”, a partir de los años 1980, técnicas de imagenología cerebral

¹⁰ Nos podemos referir aquí a la tentativa emprendida por Jaegwon Kim, de reformulación de una tesis de la “sobreviniencia fuerte”, distinta de toda perspectiva “emergentista”. Cfr. Jaegwon Kim, *el Espíritu en un mundo físico*, cap. 4, pp. 140-167 (<https://dialnet.unirioja.es/descarga/articulo/4347636.pdf>).

¹¹ Cfr. Ludwig Wittgenstein. *Observaciones sobre la filosofía de la psicología*, II, § 63: “la tercera persona debe ser identificada por observación, la primera no”.

que responden al programa de una localización intracerebral de las funciones mentales. Así, según Changeux: “en la actualidad no se duda de la existencia de las imágenes mentales. De aquí en adelante están sometidas a mediciones”.

Sin embargo, parece que la noción misma de imagen mental, identificada con una imagen “interior”, no es ni simple ni unívoca. Su carácter problemático se encontró expresamente analizado por Ludwig Wittgenstein en sus Investigaciones filosóficas. La crítica por Wittgenstein de la noción de imagen mental se despliega en particular en el marco de su estudio de los fenómenos del “ver como” y del “cambio de aspecto”, estando emparentada esta noción de aspecto con la de representación en este caso. El análisis wittgensteiniano procede a partir de esas figuras ambiguas estudiadas primero por el psicólogo Joseph Jastrow; así se puede evocar la célebre figura del “pato-conejo”, que ocupa un sitio importante en los textos de Wittgenstein consagrados al fenómeno de cambio de aspecto. Esta misma figura puede encontrarse percibida o vista, ora como representando un conejo, ora como representando un pato, pero nunca como representando simultáneamente un pato y un conejo.

El ejemplo de estas figuras desconcertantes es recuperado y ampliamente discutido por Wittgenstein, apoyando su antirreduccionismo y su rechazo de una explicación fisiológica de los fenómenos psicológicos. El análisis de tales figuras es indisociable de la crítica wittgensteiniana dirigida contra las tentativas de localización del pensamiento (desde el *Cuaderno azul*) y del rechazo de la identificación de las impresiones sensoriales con “objetos internos”¹². En el caso de tales figuras ambivalentes, la interpretación parece indisociable de la experiencia visual, lo que constituye lo propio del ver en el sentido del “ver como”.

Ahora bien, de manera notable, este mismo fenómeno “perceptivo” del cambio de aspecto se encuentra precisamente convocado por Jean-Pierre Changeux con fines de una demostración literalmente opuesta. Este fenómeno es, en efecto, encarado para apoyar su teoría de la imagen mental, concepto filosófico requerido para la teorización fisicalista de lo mental, a partir de las técnicas de imagenografía cerebral. A través de estas técnicas se podría precisamente ver la “materialidad” (sinónimo de localización) de los procesos mentales y, ante todo, de las “imágenes mentales”¹³. El ejemplo de imagen hendida retenido por Changeux no es el del pato-conejo, sino el del dibujo que representa ora un monokini, ora una copa de champaña; el fenómeno de cambio de aspecto, ligado al mecanismo de “ver como”, envuelve, según Changeux, la dimensión del sentido y del concepto. «A falta de ese que no encontré en la red, este de

¹² Esta crítica del programa de una localización del pensamiento comienza por parte de Wittgenstein desde el *Cuaderno azul* (1933-1934). Cfr. Ludwig Wittgenstein, *the Blue and the Brown Books*.

¹³ *El hombre neuronal*. Cap. Cit.

Freud viene bien al caso...)

El fenómeno del cambio de aspecto se encuentra entonces comprendido –en las antípodas del análisis wittgensteiniano– como que ilustra la idea de un parentesco (parentesco “neural”) entre los perceptos, las imágenes (en particular las “imágenes de memoria”) y los propios conceptos. Se reconoce aquí de nuevo la marca específica de una concepción empirista y asociacionista del conocimiento que hace de los conceptos los productos de una esquematización y de una abstracción progresiva de los datos sensoriales. Una tal filiación empirista se encuentra particularmente ilustrada por la referencia a Hume cuya herencia filosófica Jean-Pierre Changeux reivindica contra la tradición racionalista representada por Descartes.

Se subrayará aún que la unidad afirmada de los perceptos, de las imágenes y de los conceptos, es decir, ese parentesco estrecho sinónimo de una misma “materialidad neural”, se encuentra garantizado en *El hombre neuronal* por la categoría común de “representación”, que tendrá que ver pues con perceptos, imágenes y conceptos. Así se puede leer en el capítulo 5:

[...] hemos mostrado que al lado de las imágenes muy “concretas”, aun provistas de un contenido sensorial, existen representaciones más esquemáticas, más abstractas: los conceptos. Pero incluso si estos pueden, en ciertos casos, parecer totalmente abstractos y universales, nosotros los consideramos sin embargo como “representaciones” y los clasificamos con las imágenes en la categoría única de los objetos mentales.

Por contraste, la perspectiva wittgensteiniana evocada antes tiende a mostrar, por el análisis de esos fenómenos singulares del cambio de aspecto, la inadecuación de la noción misma de imagen mental –en tanto que supuesta imagen interna– en la medida en que ella estaría fundamentada en una confusión conceptual. Wittgenstein escribe al respecto:

[...] El concepto de “imagen interna” induce a error, pues es construido sobre el modelo del concepto de “imagen externa”; y sin embargo, los empleos de estos dos términos conceptuales no se parecen más que aquellos de “signo numérico” y “número”. [Si se quisiese llamar al número un “signo numérico ideal”, se podría suscitar una confusión análoga].¹⁴

A este respecto, nos podemos igualmente reportar a la lectura que propone Hilary Putnam, en su obra de 1999, del análisis wittgensteiniano del “ver como” y del cambio de aspecto¹⁵. Putnam adopta en esta obra una perspectiva crítica, dirigida contra el “materialismo cartesiano” contemporáneo, el cartesian cum

¹⁴ Wittgenstein. *Op. cit.*

¹⁵ Hilary Putnam. *The Threefold cord; Mind, Body, and World.*

materialism. Esta expresión designa un materialismo fisicalista, articulado a la tesis del lugar cerebral del espíritu, en tanto que este exige necesariamente un internalismo de origen –supuestamente– cartesiano, el que identifica los estados mentales con estados internos, en este caso intracerebrales. Según Putnam, el análisis wittgensteiniano del fenómeno de cambio de aspecto debería específicamente entenderse como una crítica radical de “la concepción de la interfaz”, que opera en el “materialismo cartesiano”, en este caso, en las teorías contemporáneas, reduccionistas, que identifican los fenómenos mentales con los procesos cerebrales (“brain processes”). Hilary Putnam afirma al respecto:

[...] Me parece que el célebre análisis que Wittgenstein hace de la figura del “pato-conejo” está dirigido, entre otros, contra la concepción de la interfaz. El croquis físico no es, de manera intrínseca, un dibujo ora de un conejo, ora de un pato; puede ser visto como el uno o el otro. Pero es muy difícil formar una “imagen mental” que sea ambigua en el sentido en que lo es el dibujo del pato-conejo¹⁶.

La concepción de la interfaz, siguiendo la lectura crítica de inspiración wittgensteiniana que de ella propone Hilary Putnam en *El trenzado de tres cabos*, es esta teoría de la percepción que establece que la relación del espíritu con el mundo es una relación mediada por una “interfaz” constituida de impresiones o de imágenes que residen en el espíritu. Esta concepción de la interfaz, especialmente en juego en la teoría empirista de los *sense data*, es el otro nombre del representacionalismo, y comprende dos variantes: la una dualista, que plantea el carácter inmaterial de esas impresiones o imágenes, la otra materialista. Según esta segunda versión, tales imágenes, a título de imágenes en el cerebro, son materiales¹⁷.

Ahora bien, parece que la perspectiva de Jean-Pierre Changeux –en el capítulo 5 de *El hombre neuronal*, en el que se encuentran afirmadas la materialidad (neural) de los “objetos mentales”, y el carácter paradigmático de la noción de imagen mental (cerebral) para la definición de estos objetos mentales—, implica plenamente esta concepción de la interfaz. Recordemos que esta concepción está expresamente asociada por Hilary Putnam al representacionalismo, que él designa con esa expresión polémica de cartesianism cum materialism; expresión dirigida contra el reduccionismo fisicalista y en el que “cartesianismo” es sinónimo de “internalismo”.

El postulado de un “lenguaje del pensamiento”

¹⁶ *Ibid.*, p. 45

¹⁷ Como recién traduje de G. Chazal, *Interfaces, averiguaciones sobre los mundo intermediarios...* en donde ésta es la tesis, he preferido rematar esta sección de traducciones con un paso más allá del neo-materialismo hacia el neo-mecanicismo y el neo-humanismo. Paláu.

Este mismo esquema internalista y representacionista parece aún comandar, en *El hombre neuronal*, la concepción del funcionamiento del lenguaje considerado en su relación con el “lenguaje del pensamiento”, concepción fundamentada en el postulado de una partición entre lo interno y el exterior:

[...] el lenguaje con su sistema arbitrario de signos y de símbolos, sirve de intermediario entre ese “lenguaje del pensamiento” y el mundo exterior. Sirve para traducir los estímulos o los acontecimientos en símbolos o conceptos internos, luego, a partir de los nuevos conceptos producidos, a retraducirlos en procesos externos¹⁸.

Se encuentra aquí movilizadada una teoría instrumental del lenguaje o de la significación lingüística, que coloca el primado de ese “lenguaje del pensamiento” interno, sobre el lenguaje público, “externo”.

“El hombre neuronal”, cuyo paradigma explicativo convoca un esquema reduccionista e internalista, parece entonces definirse contra lo que podríamos llamar el modelo de “el hombre simbólico”, es decir, del hombre siempre-ya inscrito en el elemento social-lingüístico.

Este lenguaje del pensamiento, lenguaje “interno”, estaría pues subyaciendo a un lenguaje más “arbitrario”, que sería el lenguaje público, el “lenguaje de las palabras”. Ese modelo del lenguaje del pensamiento converge con la teoría funcionalista del filósofo norteamericano Jerry Fodor expuesta especialmente en una obra de 1975 y titulada *the Language of Thought*, cuya perspectiva es explícitamente internalista.

Un tal modelo, el del lenguaje del pensamiento, compromete en efecto una partición conceptual crucial de lo interno y de lo externo, partición que subtiende la caracterización, en *El hombre neuronal*, del lenguaje público como “intermediario” entre el lenguaje del pensamiento y el mundo exterior. Pero la partición de lo interior y de lo exterior parece también estar en el principio de la distinción postulada por Jean-Pierre Changeux entre los dos tipos de lenguaje:

[...] El lenguaje interviene como vehículo en la comunicación de los conceptos entre individuos del grupo social. La arbitrariedad del sistema de signos implica un acoplamiento percepto-concepto de tipo “neutro”, que es el objeto de un largo aprendizaje en el curso del desarrollo [...] Al contrario, el “lenguaje del pensamiento”, permanentemente conectado con lo real, contendrá mucha menos arbitrariedad que el lenguaje de las palabras¹⁹.

Correlacionado con la representación singular de un lenguaje “natural” o de un lenguaje no arbitrario, este modelo reactiva la hipótesis, mentalista, de un

¹⁸ Changeux. *El hombre neuronal*.

¹⁹ *Ibid.*

proceso cognitivo interno, de una gramática y de una sintaxis intra-mentales que acompañan subterráneamente los enunciados lingüísticos públicos que constituirían pues su expresión segunda, la traducción externa. Este modelo del pensamiento como proceso instrumental subyacente a su significación lingüística o a su expresión pública en el elemento del lenguaje común, modelo articulado a la antítesis clásica de lo interior y de lo exterior, es precisamente el objeto de la crítica recurrente e insistente de Wittgenstein, desde las Investigaciones filosóficas hasta sus últimos escritos sobre la filosofía de la psicología²⁰.

Insistiría finalmente, en forma breve, sobre la compleja relación, o ambivalente, de Jean-Pierre Changeux, en *El hombre neuronal*, con el cognitivismo y con el modelo del espíritu-computador.

Seguramente Changeux se reclama del desarrollo de las neurociencias y, por tanto, al menos parcialmente, del cognitivismo articulado al modelo del espíritu-computador; la virtud de la “ciencia cognitiva” sería precisamente, según él, frenar al conductismo que piensa romper con la hipótesis de una “interioridad mental o cerebral” para la explicación del comportamiento humano. En efecto, la noción cognitivista-funcionalista de “codificación interna” se encuentra retomada en *El hombre neuronal*, para apoyar una hipótesis internalista según la cual “lo esencial tiene lugar dentro de la máquina”. Notemos que esta fórmula podría constituir un eslogan internalista exactamente antitética del eslogan externalista según el cual “el espíritu está en el mundo”. En el capítulo 4 podemos leer:

[...] todo comportamiento moviliza conjuntos definidos de células nerviosas y es a su nivel donde debe ser buscada la explicación de las conductas y de los comportamientos. La comparación del cerebro con una máquina cibernética, con un computador, interviene útilmente para definir esta movilización interna [...] la aprehensión del mundo exterior y la respuesta producida dependen de la organización interna de la máquina [...] Lo esencial tiene lugar dentro de la máquina, a nivel del sistema nervioso central, donde la información es transmitida según un código, analizada, y luego tratada²¹.

El autor de *El hombre neuronal*, en la medida en que propone una reactivación del paradigma de la codificación interna y de la crítica del postula conductista, parece, por consiguiente, partidario del cognitivismo. Y, sin embargo, en el capítulo 5, se aleja explícitamente de la analogía entre espíritu y computador; analogía que está en el fundamento de la teoría funcionalista-cognitivista, pero según él: “engañosa”... en razón (podemos suponer) de la tesis de la realización múltiple de lo mental. De este modo se rechaza el modelo computacional del espíritu, que compromete la tesis de la realización múltiple de lo mental,

²⁰ Cfr. en particular Ludwig Wittgenstein. *Lo Interior y lo exterior. Últimos escritos sobre la filosofía de la psicología, II*. Madrid: Tecnos, 2008.

²¹ ²⁰ Changeux. *Op. cit.*

y descalifica por consiguiente la tesis del cerebro como órgano exclusivo del pensamiento²².

De este modo podremos considerar en última instancia que Jean-Pierre Changeux retiene del funcionalismo cognitivista el postulado internalista, en este caso la hipótesis de la codificación interna en juego en el modelo del “lenguaje del pensamiento” sistematizado por Jerry Fodor, filósofo funcionalista. Por el contrario, lo que él rechaza en este mismo funcionalismo cognitivista es el basamento teórico que cuestiona la hipótesis general del lugar intracerebral del pensamiento, a saber: la comparación de la actividad mental con una secuencialidad de estados lógicos. Una tal comparación, en juego en el modelo de la máquina de Turing, supone, en efecto, como no decisivo, el tipo de sustrato o de implementación material de esta secuencialidad, y se aleja así de la representación de una sede necesariamente cerebral de lo mental.

Para concluir me detendré en dos cuestiones que se plantean hoy con la relectura de *El hombre neuronal*.

Por una parte, uno puede preguntarse si la tesis del lugar intracerebral del pensamiento, como principio de las teorías de la identidad de la que se reivindica de hecho la obra, toma plena conciencia de la crisis crónica del fisicalismo, crisis que ha dado lugar especialmente, en los años 1960, al nacimiento del cognitivismo, con la aparición del modelo computacional de espíritu-computador (máquina de Turing). Seguramente, el modelo puramente computacional del espíritu (con la primera filosofía funcionalista que lo acompaña) también entró en crisis desde hace algunos decenios, y se asiste a partir de los años 1980-1990, a una alianza entre perspectiva fundionalista y perspectiva neurobiológica como principio de la emergencia de lo que hoy se llama las “neurociencias cognitivas”. Sin embargo, este nuevo modelo neurocientífico no puede limitarse a una reactivación de las teorías clásicas de la identidad, en la medida en que no ha convertido en obsoletas las cuestiones fundamentales que están en el origen de la filosofía funcionalista, a saber: la tesis de la realización múltiple de lo mental y la cuestión de la autonomía epistemológica de este.

Por otra parte, no estamos ciertos de que el fisicalismo neurologista constituya la única respuesta a la afirmación –seguramente discutible– de una inexplicabilidad estructural de los fenómenos mentales que implicaría precisamente el fenómeno de la conciencia. La crítica wittgensteiniana de las teorías de la identidad (mental/cerebral) o, en otro registro, las teorías contemporáneas

²² *Ibid.*

del “espíritu extenso”²³, constituyen, sin duda, una vía alternativa, en ruptura con el internalismo y el representacionalismo aún comprometido en las teorías reduccionistas de lo mental.

Esta última perspectiva, que esquemáticamente se puede llamar “externalista”, busca explícitamente dar razón de la especificidad de lo mental, en este caso, de su carácter holístico, de su inscripción social, cultural, lingüística, tecnológica o, en otras palabras, de su inscripción “en el mundo”. Ella recusa simultáneamente el postulado de un esencial e inefable “misterio” de la vida mental, representada por la consciencia, y una supuesta dimensión privada irreductible del pensamiento; misterio que la perspectiva externalista se dedica precisamente a disolver, a título de confusión conceptual²⁴.

²³ Estas teorías del “espíritu extendido” (*extended mind*), llamadas igualmente teorías del “externalismo activo” plantean el rol causal determinante del entorno –físico, social, tecnológico– en la producción de los fenómenos mentales y cognitivos. Nos podemos reportar a este respecto a los trabajos del filósofo Andy Clark; ver en particular, *Being There: Putting Brain, Body and World together again; Natural-Born Cyborg: Minds, Technologies and the Future of Human Intelligence; & Supersizing the Mind. Embodiment, Action and Cognitive Extension*. El primer manifiesto de la teoría del “espíritu extendido” se puede leer en un artículo de David Chalmers & Andy Clark que data de 1998, precisamente titulado “the Extended mind”.

²⁴ ¡No podemos dejar de recomendar la lectura de François Dagognet (1992). *El cerebro ciudadela*. tr. Paláu, Facultad de Ciencias Humanas y Económicas. Área de estudios filosóficos y culturales. Universidad Nacional de Colombia. Medellín, Julio de 2002).

