

Reflexiones en torno a la filosofía de la mente

Reflections around the philosophy of the mind

Óscar Augusto García Zárate
Universidad de San Martín de Porres

Fecha de recepción: 28/09/09

Fecha de aceptación: 12/10/09

RESUMEN

El artículo aborda, desde una perspectiva que pretende ser panorámica, la problemática inscrita en el marco de las reflexiones desenvueltas por la filosofía de la mente, disciplina filosófica que ha experimentado un desarrollo vertiginoso en los últimos años. Se efectúa, así, en un contexto que, sin duda, resulta ser de actualidad, una aproximación a algunos tópicos considerados en el contexto de esta disciplina, tales como la disputa entre la psicología popular y el eliminativismo y las implicancias que se desprenden de esta última postura, las respuestas heterogéneas dadas por las perspectivas externalista e internalista en cuanto a los contenidos mentales, y los aspectos asociados a la intencionalidad y al problema de la referencia.

Palabras clave: filosofía de la mente, psicología popular, eliminativismo, contenido mental, intencionalidad, referencia.

ABSTRACT

Philosophy of mind has experienced a remarkable development in recent decades. This paper tries to summarize and explain the answers given to the main problems in this field. It reviews the disputes between externalism and internalism on mental content, as well as that between the so-called folk psychology and eliminativism, paying particular attention to the consequences of the latter. It is also concerned with intentionality and reference.

Keywords: philosophy of mind, folk psychology, eliminativism, mental content, intentionality, reference.

INTRODUCCIÓN

El problema de la mente, tal y como empieza a ser tratado por la filosofía en el siglo XX, esto es, con la exhaustividad propia de lo que ya a estas alturas constituye un ámbito disciplinar bien definido, convoca, además de los esfuerzos desplegados en este propio contexto investigativo, aquellos otros provenientes de esferas afines en cuanto a los objetivos perseguidos con respecto al conocimiento de los procesos mentales. Tenemos así, cubriendo un amplio rango de investigación, la inteligencia artificial, la psicología, la neurociencia, y en general, el despliegue de un esfuerzo multidisciplinar, en alguna medida difuso –pues incluso agrupa también algunas investigaciones aportadas desde la antropología y la lingüística–, todo lo cual conduce a plantear cada vez con mayor insistencia y plausibilidad la concreción de un quehacer novedoso: la llamada ciencia cognitiva.

Aunque la historia de la filosofía ofrece desde antaño claros testimonios de la profunda preocupación manifestada en su ámbito por los problemas relacionados con la naturaleza del pensamiento y con el enigmático mundo interior del ser humano, ya presente de manera acentuada, por mencionar algo, en un texto como el *De ánima* de Aristóteles, lo cierto es que a partir de la irrupción de Descartes es desde cuando podemos hablar de un problema más acotado, a saber, el problema de la mente humana, o, para ser más precisos, el problema de la relación mente-cuerpo. En efecto, y situándonos en un marco de reflexión formulado por Richard Rorty (1983), podríamos decir que es durante el siglo XVII cuando se inventa la mente, en el sentido de que es a partir de entonces que se sientan las bases de la concepción de la mente como aquel espacio interior en que se representan las ideas, espacio imaginado como poseedor de un estatus radicalmente distinto del cuerpo¹.

1 En el primer capítulo de la primera parte de su ya clásica obra *La filosofía y el espejo de la naturaleza*, Rorty aborda detenidamente la radical transformación que había de sufrir la concepción de la mente con el advenimiento de la modernidad y que, entre otras cosas, se manifiesta de manera sensible en el alcance que adquiere el término «idea» luego de que se le adjudicara una significación substancialmente distinta de aquella otra que poseía en la Antigüedad e incluso durante la Edad Media. En efecto, Descartes y luego Locke, por ejemplo, habían de emplear aquel término para referirse a los contenidos de la mente humana. Acerca de este radical giro dado al problema de la relación entre mente y cuerpo, dice Rorty: «La novedad estuvo en la idea de un solo espacio interior en el que eran objeto de cuasi-observación las sensaciones corporales y perceptivas («ideas confusas del sentido y la imaginación» como diría Descartes), las verdades matemáticas, las reglas morales, la idea de Dios, los talentos depresivos, y todo el resto de lo que llamamos «mental». Este escenario interior, con su observador interno, había aparecido insinuado en varios momentos del pensamiento antiguo y medieval pero nunca se había tomado en serio el tiempo suficiente como para servir de base a una problemática. Pero el siglo XVII se tomó con la seriedad necesaria para permitir que se planteara el problema del velo de las ideas (...) Demostrar que la mente podía imaginarse aparte del cuerpo era (...) un proyecto totalmente diferente del que se encuentra en la tradición procedente de Aristóteles» (RORTY, 1983: 54-56).

Sin embargo, paulatinamente, esta visión hubo de ser radicalmente cuestionada. Muestra de ello son los debates generados a partir, por ejemplo, del planteamiento del problema acerca del lenguaje privado por parte de Wittgenstein, incluso antes de la publicación de las *Investigaciones filosóficas*, en 1953. Este nuevo rumbo señalado por sus reflexiones durante la segunda etapa de su labor filosófica, la correspondiente a la desplegada por el así llamado segundo Wittgenstein, constituye un importantísimo punto de inflexión y un referente de primer orden por lo que hace al sólido posicionamiento de la temática filosófica referida a la naturaleza de los procesos mentales. Del mismo modo, ha de ser también considerado como hito importante en la trayectoria seguida por esta temática la publicación, poco antes, en 1949, por parte Gilbert Ryle, de otra obra capital: *El concepto de lo mental*. Pero, además, de ningún modo podría soslayarse una cuestión clave en todo esto. En efecto, de la mano con este despliegue auroral del horizonte reflexivo que la pregunta por los asuntos relativos a la mente empezaba a configurar marchaban, asimismo, los desconcertantes avances que se iban logrando, en un principio, en el campo de las investigaciones referidas a la informática y a la inteligencia artificial, y, un poco más tarde, los no menos cruciales descubrimientos hechos en el terreno de la neurociencia. Todo ello, hace apenas seis o siete décadas. Así, pues, es de reciente data el notable impulso interdisciplinar que ha sido dado a la reflexión en torno a la relación entre la mente y el cerebro, tal y como se desenvuelve en las lindes de la filosofía, esto es, adoptando diversos matices y énfasis, y a través de acuerdos y desacuerdos, marchas y contramarchas, sin que ello signifique dejar de seguir muy de cerca los aportes de las ciencias.

«Un cálculo lógico de las ideas inmanentes en la actividad nerviosa», artículo escrito por Warren McCulloch y Walter Pitts, y publicado en 1943 en el *Bulletin of Mathematical Biophysics*, muy probablemente constituye una de las más claras muestras del modo cómo empezó a asumirse la investigación de la denominada inteligencia artificial. Sin duda este artículo posee el carácter de un documento fundacional en este tema y dio origen al surgimiento de dos tendencias en el marco de las investigaciones acerca de la inteligencia artificial y la mente humana: por un lado, la llamada variante tradicional, y, por otro, el conexionismo (BODEN, 1994: 10-11).

Al abrigo de este inicial impulso, asimismo, y luego de haberse desarrollado, en septiembre de 1948, un simposio en el Instituto de Tecnología de California acerca de «Los mecanismos cerebrales de la conducta», auspiciado por la Fundación Hixson –y en donde al lado de oradores como John von Neumann y Karl Lashley se encontraba el mismo McCulloch–, el año 1956 tuvo lugar en el *campus* del Dartmouth College,

un encuentro académico al que asistió un grupo selecto de científicos que realizaban investigaciones en el terreno de la inteligencia artificial. Cuatro de ellos –John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Elwood Shannon– fueron los encargados de redactar la propuesta hecha ante la fundación Rockefeller para la realización de dicho evento. Es célebre ya el tenor de su pasaje inicial. En él se plantea el principal propósito perseguido por los investigadores:

Proponemos que se lleve a cabo una revisión de la inteligencia artificial de dos meses y diez hombres en el verano de 1956 en el Dartmouth College de Hanover, New Hampshire. La revisión es para seguir adelante sobre la conjetura de que cualquier aspecto del aprendizaje o cualquier otro rasgo de inteligencia puede en principio ser descrita de forma tan precisa que puede hacerse que una máquina lo simule. (MC-CORDUCK, 1991:105)

Curiosamente, ese mismo año, se celebraba en el Instituto de Tecnología de Massachusetts una reunión que bajo la denominación de Simposio sobre Teoría de la Información se proponía elaborar los lineamientos que había de seguir una ciencia de la mente, siguiendo de cerca e incorporando los conocimientos forjados en los terrenos de la informática, y que muchos consideran que se trata del momento en que oficialmente entra en la escena académica la denominada ciencia cognitiva. Entre los más significativos aportes que se dieron en este evento pueden mencionarse dos. En principio, los correspondientes a la ponencia de Allen Newell y Herbert Simon que describía el curso seguido por «La máquina lógica» en la primera demostración completa de un teorema. Y, además, también las ideas acerca del enfoque transformacional de la gramática formuladas por el destacado lingüista Noam Chomsky (GARDNER, 1987: 44).

Éstas son, a vuelo de pluma, las coordenadas históricas por lo que hace al curso heterogéneo seguido por las investigaciones que diríase en un esfuerzo prometeico buscan echar luz sobre aquellos aspectos de la mente humana que parecen obstinarse en permanecer ocultos a la curiosidad humana. El trazado modesto de este panorama a través del cual apreciamos el abigarrado encuentro de diversas perspectivas de investigación, que, no obstante, articulan un mismo interés, a saber, el afán de desentrañar el funcionamiento y, si cabe, la naturaleza de la mente, nos permitirá abordar prospectivamente algunos de los temas que constituyen el ámbito de investigación de la filosofía de la mente.

ELIMINATIVISMO Y PSICOLOGÍA POPULAR

Un problema que, según nos parece, merece ser considerado en primer lugar es el relativo a la psicología popular y los embates que ha soportado por parte del eliminativismo, programa éste encabezado por Paul y Patricia Churchland.² En principio, ¿qué podemos decir acerca de la psicología popular? Veamos.

A diario desenvolvemos una gran cantidad y variedad de acciones que son desencadenadas por ciertos estados mentales. Miedos, deseos y creencias conforman, entre otros estados intencionales, el caleidoscopio interior que constituye el mundo mental del ser humano. A partir de aquellos estados mentales articulamos una serie de explicaciones y predicciones que sin estar necesariamente basadas en conceptos forjados en los términos propios de una rigurosa investigación científica, sirven, sin embargo, para dar cuenta de las acciones de las demás personas y, claro, también de las nuestras. Así, pues, la psicología popular es la base explicativa que, articulada en torno a instancias intencionales, subyace a la visión cotidiana del sentido común. Precisamente por esto último, este tipo de psicología –molesto obstáculo para los partidarios de la neurofilosofía³ tanto como para el eliminativismo radical– también es denominada psicología del sentido común.

- 2 Habría, sin embargo, que hacer una necesaria precisión. Aun cuando Patricia y Paul Churchland asumen una actitud de frontal cuestionamiento frente a la psicología popular, existe, con todo, una marcada diferencia entre el eliminativismo intencional propuesto por Paul Churchland y aquel otro formulado por Patricia Smith Churchland, que, a diferencia de aquél, asume un carácter sentencial. El modelo sentencial, que es cuestionado por Patricia Churchland, postula básicamente la existencia de una estructura interna computacional de naturaleza lingüística como la única posibilidad de justificar la conducta de los sistemas intencionales. De modo conciso, la posición de Patricia Churchland frente a este modelo podría quedar expuesto en los siguientes términos: «Si la Psicología Popular es una teoría verdadera, entonces, en efecto, la mente es una obra de ingeniería lingüística tal y como la hipótesis fodoriana del lenguaje del pensamiento lo plantea. Pero el paradigma lingüístico de procesamiento de información es falso, luego la Psicología Popular también lo es» (TORIBIO, 1995: 254). Por su parte, Paul Churchland presenta en los siguientes términos los alcances antiintencionalistas de su materialismo eliminativista: «Hace veinte años, las emociones, los *qualia*, y los «sentimientos puros» se postulaban como los principales escollos del programa materialista. Con estas barreras derribadas, el lugar de oposición se ha desplazado. Ahora éste es el reino de lo intencional, el reino de la actitud proposicional, que se presenta la mayor parte de las veces como si fuera irreductible e imposible de eliminar en favor de cualquier cosa que se sitúe dentro de un marco materialista» (CHURCHLAND, 1981: 67-68).
- 3 Fue Patricia Smith Churchland quien en 1986 publicó *Neurophilosophy*, una obra en que además de exponer un detallado cuadro del avance impresionante que se ha logrado en materia neurocientífica, expresa su pleno y entusiasta convencimiento de que se cuenta ya, por ello mismo, con herramientas científicas que han de hacer posible consolidar lo que la autora llama en algún momento una ciencia unificada de la mente. Haciendo alusión al nuevo panorama que ella avista en relación con el estudio de la mente, dice: «La convicción que sostiene este libro es que las estrategias *top-down* (que caracterizan a la filosofía, a la psicología cognitiva y a la investigación en inteligencia artificial) y las estrategias *bottom-up* (que caracterizan a las neurociencias) para

Frente a esto, el eliminativismo sostiene que el recurrir a estados intencionales para explicar la conducta de las personas constituye una herramienta conceptual que no posee contenido alguno. Las actitudes intencionales, dicen los abanderados de esta corriente, no poseen una condición causal, pues

Nada que no sea físico puede actuar como causa. Luego, la apelación a tales estados intencionales como causalmente responsables de nuestra conducta es una apelación improcedente. (TORIBIO, 1995: 246)

Según esto, el eliminativismo pasa a postular, entre otras cosas, que las propiedades a las que se denominan «psicológicas» no son de ningún modo coextensivas con las propiedades de naturaleza física. Por otra parte, sostiene que en la medida que los estados neurofisiológicos son estados internos del individuo, los paulatinos descubrimientos que se vienen llevando a cabo en el terreno de la neurociencia proveerán el nivel de conocimientos necesario para dar cuenta en claros y estrictos términos explicativos de la dinámica profunda de los procesos cognitivos. Sobre la base de esta concepción, los teóricos del eliminativismo aseguran que es posible dejar de postular la existencia de una entidad que permitía explicar un fenómeno –en este caso se trataría del espurio armazón explicativo de tipo intencional con que cuenta la psicología popular– en el preciso momento en que se halle una manera novedosa de explicarlo en substitución de aquélla –también en este caso se trataría de los descubrimientos en materia de neurofisiología–, de modo tal que, en el caso que nos ocupa, la descripción y explicación de un fenómeno mental ha de excluir el uso de un lenguaje con contenido psicológico. (RABOSI, 2000)

resolver los misterios de la función mente-cerebro no han de ser desenvueltas en un gélido aislamiento. Lo que se entrevé, por el contrario es una rica interacción entre ambas, de la que se pueda esperar que provoque una fecunda y paralela evolución de teorías, modelos y métodos que se informen, corrijan e inspiren mutuamente» (CHURCHLAND, 1986: 3). Y en otro lugar, refiriéndose también a la neurofilosofía, expresa lo siguiente: «En evolución paralela con la neurociencia, la ciencia cognitiva ha investigado el alcance de funciones de gran escala tales como la atención, la memoria, la percepción y el razonamiento tanto en el adulto cuanto en el desarrollo del niño. Adicionalmente, las ideas computacionales al vincular fenómenos *cognitivos* de gran escala con fenómenos neurales de pequeña escala han abierto la puerta a una integración de la neurociencia, la ciencia cognitiva y la filosofía en un marco teórico comprehensivo. (...)

¿Qué es la neurofilosofía y cuál es su rol? (...) La neurofilosofía surge del reconocimiento de que después de mucho, finalmente, las ciencias del cerebro y su tecnología están suficientemente desarrolladas de modo que se puede llegar a alcanzar un real progreso en la comprensión de la relación mente-cerebro. Más aún, ella predice que la filosofía de la mente conducida sin ninguna comprensión de las neuronas y el cerebro sea probablemente estéril. Como resultado de ello, la neurofilosofía centra su atención en problemas que se encuentran en la intersección de una neurociencia que va echando raíces y una filosofía en declive» (CHURCHLAND, 2002: 2-3)

Un aspecto puesto especialmente de relieve por la crítica eliminativista es aquél referido a la presunta condición de teoría empírica que ostentaría la psicología popular. Esto no conduce a otra cosa sino a considerar que una explicación dada en términos de deseos, creencias o expectativas, esto es, una explicación que asume la forma de algunas de estas modalidades del sentido común, o, en otros términos, cada una de las preferencias que dan cuenta de una actitud proposicional, funciona como una instancia particular de una generalización de carácter nomológico. Así, cuando nos referimos al hecho de que Juan estudia con ahínco *diciendo* que la razón de ello es que Juan *desea* ingresar a la universidad, nuestra explicación vendría a ser una instancia particular de una generalización que adoptaría la estructura siguiente: Para cualquier individuo x , cualquier deseo p y cualquier creencia q , si x (Juan) desea p (ingresar a la universidad) y x cree que si q entonces p (Juan cree que si estudia con ahínco ingresará a la universidad), entonces, a menos que otras creencias o deseos impidan su realización, x (Juan) realiza q (estudia con ahínco).⁴

Esta presunta estructura teórica nomológica es puesta en cuestión por el eliminativismo. En la medida en que la psicología apela a instancias «mitológicas» tales como las actitudes proposicionales con el objeto de explicar los procesos mentales, y teniendo en cuenta que se trata, como se ha dicho, de una teoría, toda vez que recurre a generalizaciones del tipo a que se ha hecho referencia, pues entonces, se concluye, la psicología popular no es sino una teoría empírica del todo falsa, y, por ello mismo, debe ser abandonada. No se trata, pues, de buscar reconducir las nociones intencionales de este ámbito a un terreno científico a través, digamos, de leyes puente –en este caso estaríamos hablando de la aplicación de una estrategia reduccionista–,⁵ sino, más bien, y como lo pretende, Paul Churchland, de reemplazar completamente las explicaciones intencionales por un vocabulario articulado sobre la base de conceptos provenientes de la neurociencia:

4 La estructura de esta generalización la he tomado de TORIBIO, 1995: 245-272. Sin embargo, a fin de incorporar alguna dosis de originalidad al ejemplo he considerado oportuno variar su contenido.

5 En este punto, justamente, es donde se separan eliminativismo y materialismo reduccionista. En efecto, mientras aquél rechaza categóricamente que pueda efectuarse una reducción de los principios y de la ontología que implica el empleo del lenguaje intencional a términos provenientes de las investigaciones en neurofisiología, el materialismo reduccionista –que patrocina la denominada tesis de la identidad– estima viable una reducción interteórica de las nociones intencionales de la psicología popular, de modo tal que una vez hecho esto, todo proceso mental que era explicado apelando a dichas nociones es reconducido al interior de un marco teórico vertebrado tomando como eje el lenguaje y las conquistas científicas de la neurociencia (CHURCHLAND 1999: 50-63).

El materialismo eliminativista es la tesis de que nuestra concepción de sentido común acerca de los fenómenos psicológicos constituye una teoría radicalmente falsa, una teoría tan fundamentalmente defectiva que tanto sus principios cuanto su ontología serán, antes que apaciblemente reducidos, eventualmente substituidos por el cabal desarrollo de la neurociencia. Nuestro mutuo entendimiento e incluso nuestra propia introspección podrían entonces ser reconstituidos dentro del marco conceptual de una neurociencia adecuadamente desarrollada, una teoría que de lejos podríamos esperar que sea más potente que la psicología del sentido común que aquélla reemplazaría, y más substancialmente integrada en el contexto de la ciencia física. (CHURCHLAND, 1981: 67)

Haciendo frente a este intento por acabar de una buena vez con la psicología popular y su entramado intencional de explicaciones, se encuentra Jerry Fodor, uno de sus más entusiastas defensores. En el contexto mayor de lo que él denomina Teoría Representacional de la Mente, Fodor asume la reivindicación de la psicología popular, al punto de aspirar a convertirla en una ciencia respetable. Haciendo gala no sólo de solidez argumentativa, sino también de una elegancia expositiva a menudo cautivante, Fodor lleva a término una cerrada defensa de la psicología popular, a la que se refiere como «la vieja psicología de sentido común de las creencias y los deseos» (FODOR, 1994: 20). Esta urdimbre de supuestos entretejida con deseos y creencias y conectada con acciones es la que se despliega habitualmente en medio de predicciones que se hacen casi sin notarlo, y de acuerdo a las cuales las más diversas actividades adquieren un orden y un sentido: la defenestrada psicología popular está en todas partes y en todo momento, mostrando con esto que, aunque casi no sea advertida, su presencia es notablemente importante.

[La psicología popular] es la que nos dice, por ejemplo, cómo inferir las intenciones de las personas a partir de los sonidos que pronuncian (si alguien emite la forma de palabras «llegaré al aeropuerto en el vuelo de las tres de la tarde», entonces, *ceteris paribus*, intentará llegar al aeropuerto en el vuelo de las tres de la tarde). También nos dice cómo inferir el comportamiento de las persona a partir de sus intenciones (si alguien tiene la intención de llegar al aeropuerto en el vuelo de las tres, entonces, *ceteris paribus*, se comportará de una forma tal que lo que sucederá será que llegue a ese lugar a esa hora, dejando al margen fallos mecánicos y actos divinos. (FODOR, 1994: 20)

LAS IMPLICANCIAS DEL ELIMINATIVISMO

Luego de esta breve exposición de lo que podría considerarse el estado de la cuestión en lo que respecta al debate entre eliminativistas y aquellos que, como en el caso de Jerry Fodor, otorgan un lugar especial a la intencionalidad en la explicación de los procesos mentales, sería oportuno referirnos a lo que creemos se presenta como una consecuencia nefasta que se desprende del eliminativismo: la concomitante, aunque no explicitada, eliminación de la noción de persona. Pasemos a considerar esta cuestión.

Algún autor (SANFÉLIX, 2000) ha trazado con acierto y agudeza la tortuosa ruta que ha seguido la noción de persona en el contexto de la filosofía de la mente. De ser convertida en foco de atención y puesta en un lugar de expectativa en autores como Wittgenstein y Strawson, dicha noción fue perdiendo protagonismo paulatinamente, luego de la eclosión de la llamada revolución cognitiva, hasta verse expuesta a la total exclusión en un marco teórico como el formulado por el eliminativismo.

Así pues, los enfoques cognitivos desde la época de aquella revolución han abordado el problema de la relación mente-cuerpo dentro de un marco que puede ser llamado naturalista, y a partir de ello dieron origen básicamente a dos líneas de investigación, dejando ambas siempre de lado la noción de persona. Tales orientaciones fueron el funcionalismo computacional y el eliminativismo, y es éste, como se hace evidente, el que concentra interés en este punto.

El eliminativismo, como tuvimos ocasión de ver, convierte en el centro de su atención más que los avances en materia de inteligencia artificial, el papel –que estiman crucial aquellos que están entre sus filas– cumplido por los descubrimientos efectuados por la neurociencia en lo que atañe a las respuestas que, sobre esta base, se podrían alcanzar en referencia al problema de la mente humana. Como también se ha indicado, el propósito perseguido por esta corriente consiste en erradicar las actitudes proposicionales encuadradas en el contexto explicativo de la psicología popular, pues, aducen los eliminativistas, un lenguaje articulado en torno a aquéllas, y por esto mismo típicamente intencional, no es susceptible de ser reducido a los principios y proposiciones que conforman el marco teórico de la neurociencia. Y ello por una sencilla razón: la psicología popular constituye una teoría empírica radicalmente falsa.

Pero he aquí que surge una situación por lo demás alarmante: si el eliminativismo obra en consecuencia, el desenlace de su proceder será simple y llanamente la concomitante eliminación de la noción de persona,

con todo lo que ello implica. Pues, en efecto, ése sería la lamentable consecuencia de la cruzada científicista emprendida contra el núcleo de la psicología popular, es decir, la intencionalidad, ya que ésta es considerada no sin razón un componente básico de la personalidad. El panorama que nos sería dado avistar luego de producirse un despojo de este tipo sería desolador ciertamente:

Tomémonos en serio, siquiera sea por un momento, la posibilidad de una situación en la que el lenguaje intencional ha desaparecido. En un contexto semejante los seres humanos carecerían de una biografía, o al menos de una biografía que guardara semejanza alguna con nuestro concepto de tal. Pues la biografía que nos identifica personalmente es una biografía que narramos utilizando el lenguaje intencional. Lo que creímos y descreímos, aquello en lo que seguimos creyendo, los deseos ya muertos y los que todavía mantenemos vivos, los satisfechos y los frustrados, nuestros propósitos de enmienda o nuestra firme intención de perseverar... he aquí lo que nos define ante los ojos de los demás ya ante los nuestros propios. (SANFÉLIX, 2000: 337)

Y aún hay más. El riesgo que lleva aparejado una estrategia de este tipo –aun cuando el eliminativismo continúe hablando de personas, si bien, y adoptando el fraseo rortyano, de «personas sin mente»– el riesgo, decíamos, se hace patente al momento de reparar en que estamos frente a la posibilidad de una eventual cosificación de las relaciones humanas. En efecto, si la conducta humana es estudiada y evaluada en fríos términos de estados neurofisiológicos, luego de operarse una hipotética substitución del componente intencional inherente a ella, y con esto, anularse la dimensión personal, por parte de una neurociencia notablemente desarrollada, se insinúa la abominable posibilidad de instaurar estrategias más efectivas de control y manipulación social como producto del poder que se alcanzaría al conocerse los mecanismos que permitirían manejar a un nivel neurotecnológico los estados físico-químicos del cerebro. Un horizonte sombrío como éste, aun cuando parezca formar parte de las páginas de alguna distopía filosófica –piénsese en *Un mundo feliz*, de Huxley, o en *1984*, de Orwell–, nos permite, eso creemos, tomar conciencia de las desconcertantes ideas, que incubándose inadvertidamente en doctrinas al uso, podrían llegar a convertirse al cabo en una triste y perversa realidad.

Pero dejemos las cosas en este punto. El simple propósito en relación con esta temática era dejar bosquejado lo principal de esta incisiva objeción a las implicancias del eliminativismo.

CONTENIDO MENTAL RESTRINGIDO Y CONTENIDO MENTAL AMPLIO

Pasando a tratar ahora otra cuestión, valdría la pena aproximarnos al debate particularmente interesante suscitado por el problema concerniente a los estados mentales y el contenido a ellos correspondientes. Algunos apuntes en torno al célebre experimento mental ideado por Hilary Putnam (1975), aquél conocido como el de la Tierra Gemela, resultaría apropiado para efectuar un bosquejo de esta problemática.

Putnam, a través de este ejemplo de ciencia ficción, como él lo denomina, se propone mostrar los problemas que entraña el uso del término «significado», pues estima que éste descansa en una falsa teoría que, a su vez, se encuentra articulada en torno a dos asunciones que no están convenientemente justificadas. Así, señala Putnam, sobre la base de estos dos supuestos, se afirma, por una parte, que conocer el significado de un término es estar en un determinado estado mental, y, por otra, se sostiene que el significado, en términos de intensión, es el que fija la extensión. Al momento de discutir, estas presunciones, Putnam aborda lo que él denomina solipsismo metodológico, es decir, la concepción acerca de que ningún estado mental presupone la existencia de otro individuo a no ser sólo la de aquél a quien se adscribe un estado mental. A este tipo de estados los llama estados mentales en sentido restringido. Frente a éstos, Putnam propondrá los estados mentales en sentido amplio, los cuales, a diferencia de aquellos otros, suponen lo que en otro texto llamará «contribución del entorno» (PUTNAM, 1995: 63), vale decir, la determinación del significado supone la participación de condiciones externas. Dicho de otro modo, y en palabras de Putnam: «¡Los significados no están en la cabeza!» (PUTNAM, 1975: 227).

Dicho experimento, y según lo que venimos diciendo, apunta a poner de relieve el rol decisivo que desempeña el entorno físico, o, en otros términos, las condiciones externas, con vistas a determinar el significado asociado a una determinada representación mental. Como se sabe, Putnam introduce en esta historia dos personajes que son gemelos idénticos y que viven uno en nuestro planeta y el otro en un planeta muy semejante al nuestro que, por ello, mismo, recibe el nombre de Tierra Gemela. Respectivamente, los nombres de estos personajes son Óscar₁ y Óscar₂. Al ser planetas muy semejantes, en la superficie de ambos hay agua. Sin embargo, aquello a que se refieren estos dos personajes cuando señalan el agua no podría de ningún modo ser lo mismo. Putnam al respecto señala que incluso en el caso de que Óscar₁ y Óscar₂ se encontraran en una época anterior al descubrimiento de la composición molecular del agua, digamos

en 1750 (que en el caso de la Tierra había de ser determinada como H_2O ; y en el de la Tierra Gemela como XYZ), la situación, en términos de significado, no variaría, Óscar₁ y su doble seguirían refiriéndose a dos cosas distintas, pues aun cuando se encontraran en estados mentales idénticos, y a la ciencia de sus respectivos mundos les tomara cincuenta años más descubrir la composición del agua, la extensión del término empleada para designarla siempre sería distinto: «agua» para el terrícola Óscar significaría H_2O , en tanto que para el habitante de la Tierra Gemela sería XYZ. Insistamos en esto: el significado de un término, según Putnam, no queda fijado por los estados mentales de las personas (si fuera así, Óscar₁ y Óscar₂ estarían en posesión del mismo significado), sino por la referencia. En el caso de Óscar₁ y su sosias idéntico, no es el caso que el significado haya variado en el transcurso de alrededor de cincuenta años (el período aproximado que, desde el año 1750, le tomaría a la ciencia descubrir la composición química del agua), sino son las *condiciones* externas –o como también lo dice Putnam, la relación teórica, esto es, la relación de mismidad (*sameness*) que este término de clase natural (aquello a que llamamos agua) establece con aquel tipo de cosas con que se identifica– las que en última instancia han sufrido una modificación, condiciones que por lo demás siempre estarán expuestas a ser modificadas por eventuales descubrimientos posteriores (PUTNAM, 1975: 224-225).

El planteamiento de esta temática por parte de Putnam se considera la fuente que genera la explicitación del debate a que nos referimos (TORIBIO, 2000: 233-234; ACERO, 1995: 182-183). Putnam, como se recordará, conspicuo representante temprano del funcionalismo,⁶ al proponer una explicación del significado que sea articulada, como hemos visto, tomando como eje el llamado contenido mental amplio, se sitúa de esta manera en las filas del externalismo, adversario teórico del internalismo asumido, entre otros, por autores como Jerry Fodor, y que, en consecuencia, se muestran a favor de la explicación de la problemática relativa a la intencionalidad en términos de lo que se denomina contenido mental restringido.

6 Vendría bien señalar, sin embargo, que Putnam tomó distancia de su original postura, y sometió a un balance crítico los principales supuestos del funcionalismo por él propuesto en la década de los sesenta. Refiriéndose al propósito que persigue en la obra en que manifiesta esta actitud, dice Putnam en la introducción de aquella: «(...) me propongo demostrar que los argumentos a favor de la concepción computacional, esto es, los mismos argumentos que yo solía esgrimir para demostrar que es incorrecto identificar los estados mentales con estados físico-químicos, pueden generalizarse y extenderse de modo que resulte también incorrecta la identificación directa de los estados mentales con estados *funcionales*, es decir, con estados caracterizados computacionalmente» (PUTNAM, 1995: 15).

INTENCIONALIDAD Y SIGNIFICADO

En este punto no estará de más recordar otra historia de ciencia ficción elaborada por Putnam para explorar la noción de referencia. En este caso se trata de la historia a través de la cual Putnam invita al lector a imaginar que, en realidad, los seres humanos y, en general, todo ser sintiente, son cerebros en una cubeta conectados a un sofisticado mecanismo que induce en ellos, a través de una complejísima red de *inputs*, la ilusión de la existencia de un mundo externo en la que son personas de carne y hueso. Estimo pertinente incorporarlo en este contexto porque nos permitirá entrar a abordar luego otro *Gedankenexperiment*, que resultará pertinente traer a colación, aquél debido a John Searle. Pues hay, a nuestro juicio, una relación entre estos dos puntos de vista: tanto Searle como Putnam niegan a toda costa que el nivel de manipulación simbólica en una máquina implique necesariamente un nivel de comprensión de significados y, con esto, de intencionalidad.

Si bien el experimento en cuestión, tiene como trasfondo un objetivo distinto, esto es, demostrar, arribando a un desenlace en que el argumento se autorrefuta, que si pensamos que somos cerebros en una cubeta, en realidad, no somos cerebros en una cubeta, no obstante, uno de los aspectos principales en el contexto de la argumentación es la exploración de la noción de referencia. Para los propósitos de la presente exposición bastará tomar en cuenta sólo esto último. Veamos.

Putnam, en principio, presenta el fantástico si bien lógicamente posible caso de una hormiga que desplazándose parsimoniosamente sobre la arena trazara una línea de contornos caprichosos que a la postre configurara una caricatura de Winston Churchill. Frente a esto, la pregunta que se podría formular, sería la siguiente, ¿dicha caricatura se «refiere» a Churchill? La respuesta es un rotundo no. Luego se verá cómo se incorpora esta respuesta en el marco argumentativo global que formula Putnam. Digamos entretanto que a partir de esto, y pasando a otro plano de la argumentación, Putnam apela al Test de Turing, aquella prueba ideada por el célebre matemático británico para determinar los márgenes dentro de los cuales resultaría lícito adjudicarle a una computadora procesos de pensamiento semejantes a los que despliega el ser humano, y sobre esta base señalar en la máquina la presencia de conciencia, luego de mantener una conversación con ella, sin estar al tanto, claro está, de que efectivamente se trata de una máquina de este tipo, de modo que no podamos distinguir sus respuestas a nuestras preguntas de aquellas otras que daría una persona. Putnam efectúa una ligera variante al test, pues en su caso, él lo tomará en cuenta para determinar si es posible que

una computadora logre referirse a algo, es decir, desenvolver procesos de tipo intencional sobre la base de los cuales sea capaz de elaborar representaciones que hablen del mundo, tal y como lo haría un ser humano, de modo tal que la posibilidad de hacer una distinción tajante entre el modo de referirse a objetos por parte de la máquina y el correspondiente a una persona –aquello a lo Putnam llama la referencia compartida– quede de plano cancelada. Putnam denomina esta variante de la prueba Test de Turing para la Referencia.

Si imaginamos, dice Putnam, que nos enfrascamos en una conversación con un interlocutor, sin saber que se trata de una máquina, al punto de que en el curso de este juego de la imitación, dicha máquina logra engañarnos, y como producto de ello llegáramos a la errónea conclusión de que se trata de una persona, de lo cual también se desprendería que esa máquina efectivamente ha logrado referirse a los hechos del mundo en los exactos términos en que lo hacen las personas, lo cierto con respecto a todo esto es que, en verdad, aquella máquina no habría llegado a instalarse en un plano representativo, referencial. Y esto por una sencilla razón: la manipulación de símbolos por sí mismo no es garantía de que se produzca un uso referencial de esos símbolos. Y es en esta medida que resulta legítimo decir que la capacidad referencial de una máquina con esas características es la misma que la que por cuestiones de azar ha puesto en juego la fantástica hormiguita dibujante que se mencionaba líneas arriba: es decir, dicha capacidad es inexistente. Putnam se refiere así a las supuestas capacidades referenciales espuriamente atribuidas a una computadora que haya aprobado con éxito el Test de Turing para la referencia:

Me parece evidente que ni podemos ni debemos atribuir uso referencial a tal ingenio mecánico. Es cierto que la máquina puede platicar maravillosamente acerca del paisaje de Nueva Inglaterra, por ejemplo. Pero si tuviese enfrente una manzana o una vaca, no podría reconocerlas.

Tenemos un artilugio para producir oraciones en respuesta a oraciones. Pero estas oraciones no están conectadas con el mundo real. Si *acoplásemos* dos de estas máquinas y les dejásemos disputar el juego de la imitación entre sí, *continuarían «engañándose» una a otra eternamente, aun cuando el resto del mundo desapareciese.* (PUTNAM, 1988: 23-24)

Y luego, añade algo que, según lo estimamos, vincula estas conclusiones a las que presenta Searle en el experimento del cual dentro de poco nos ocuparemos:

(...) no hay ninguna razón para considerar el discurso de la máquina (...) como algo más que un juego sintáctico. Un juego sintáctico que se *parece* a un discurso inteligente, a buen seguro. Pero en la misma medida (y no en mayor) que la curva de la hormiga se parece a una mordaz caricatura. (PUTNAM, 1988: 24)

Parece ser legítimo a este respecto indicar, como ya lo habíamos dejado apuntado antes, que se puede detectar una coincidencia de pareceres dada entre la perspectiva de Putnam y las ideas que animan la concepción de Searle respecto de lo que cabría denominar «estupidez semántica» en relación con los procesos de manipulación simbólica realizados por computadoras. Pues, en efecto, creemos que es posible encontrar un punto de intersección entre la tesis que Searle defiende a través del experimento mental de la habitación china, a saber, que las máquinas no poseen una dimensión semántica que les permita comprender significados –y a través de la cual establece una polémica también con los principales supuestos asumidos en el marco de lo que él llama inteligencia artificial fuerte–, y aquella otra defendida por Putnam, que acabamos de exponer, en el sentido de que no es posible atribuirle capacidad referencial a una computadora aun cuando ésta desenvuelva impecables procesos de manipulación sintáctica.

En lo que atañe al célebre experimento mental conocido como La habitación china, y a través del cual Searle (SEARLE, 1994) también recusa la idea procedente de Turing acerca de que una máquina, esto es, una computadora, podría ser capaz de desplegar un tipo de procesos que podrían ser considerados como pensamiento, la conclusión a la que se arriba resulta ser categórica: las máquinas de ese tipo sólo procesan información en términos sintácticos, del mismo modo como el hombre de su experimento manipula de manera eficaz ideogramas chinos en la habitación en la que se encuentra encerrado, lo que le permite generar un proceso que es considerado eficaz por las personas que se encuentran fuera de la habitación, y esto aun cuando no comprenda absolutamente nada de ese idioma. A diferencia de la ejecución de una tarea como ésta, la comprensión implica la puesta en marcha de un proceso semántico, proceso que las máquinas no están capacitadas para cumplir.

En relación con lo dicho, Searle adopta una postura que sin dejar de reconocer el origen causal de la bioquímica cerebral, otorga, del mismo modo, su lugar a la intencionalidad, y de acuerdo a ello, defiende la tesis según la cual los procesos mentales dependen de procesos cerebrales,

si bien son irreductibles a éstos⁷. A esta postura algunos autores la han denominado emergentismo (MARTÍNEZ, 2002). Si bien, dice Searle, la conciencia es un fenómeno que puede ser remitido en su origen a procesos neurológicos, no obstante, aquélla no se agota en éstos. Por la manera en que Searle trata el tema concerniente a la relación mente-cerebro, se ha dicho que es él quien ha reintroducido el tema de la conciencia en medio del debate en torno a esta problemática (GARCÍA, 2006; GONZÁLEZ-CASTÁN, 1992). Searle estima que el estudio de la conciencia puede llevarse a término empleando una metodología científica (SEARLE 1998), y, según esto, considerando la base neurofisiológica de aquélla, pero sin que ello suponga negar, tal y como lo hace el eliminativismo, la intencionalidad como fenómeno relevante al momento de explicar la naturaleza de los procesos mentales. Asimismo, su desacuerdo es frontal con respecto a la tesis esgrimida por los teóricos de lo que Searle identifica como expresión, algo que ya mencionamos, de la inteligencia artificial fuerte, y merced a la cual se busca reproducir y explicar el fenómeno mental a través del diseño de programas de cómputo, pues, sostienen sus teóricos, éstos poseen la particularidad de poder ser ejecutados por cualquier máquina adecuadamente construida; en buena cuenta, tal y como lo hace ver Searle, esta perspectiva supone un dualismo, que aunque no es típicamente cartesiano, de todos modos postula la independencia del programa (la instancia que guarda semejanza con la mente, según los partidarios de la inteligencia artificial fuerte) en relación con su concreta ejemplificación por parte de cualquier «cerebro» electrónico. Este dualismo es rechazado por Searle, pues él sostiene, más bien, que si bien el cerebro humano puede ser considerado una máquina, se trata de un tipo de máquina que se relaciona intrínsecamente con los fenómenos mentales. En el campo de la inteligencia artificial, por el contrario, no se postula ninguna relación intrínseca entre los programas y los «cerebros» que los ponen en marcha. Las únicas máquinas que al igual que el cerebro humano podrían ser asiento de procesos de pensamiento y conciencia serían aquellas que contaran con los mismos poderes causales con que cuenta el cerebro, de modo tal que mal podría decirse, si se asume el marco emergentista, que los procesos mentales (que pretenden ser reproducidos en alguna escala por los teóricos de la inteligencia artificial a través de programas) son independientes de su base física (SEARLE 1994: 102).

7 Su postura al respecto queda claramente expresada en los siguientes términos: «Yo entiendo inglés y poseo otras formas de intencionalidad, no porque soy el ejemplo concreto de un programa de cómputo (...), sino porque, hasta donde sabemos, soy un cierto tipo de organismo, con una cierta estructura biológica (es decir, cierta estructura química y física), la cual, en ciertas condiciones, es causalmente capaz de producir percepción, acción, comprensión, aprendizaje y otros fenómenos intencionales» (SEARLE 1994: 97).

Y así pues, y volviendo al punto relacionado con la intencionalidad, esto es, la capacidad referencial, propia de la mente humana a juicio de Searle, el filósofo norteamericano niega rotundamente que una máquina de cómputo pueda duplicar los procesos mentales humanos en términos de intencionalidad. Y así, ante la pregunta acerca de si la ejemplificación concreta de un programa a través de una de estas máquinas no humanas podría ser considerada, por sí misma, condición suficiente para el surgimiento de la comprensión, Searle contesta con un categórico no. Pues él considera que

(...) las manipulaciones de símbolos formales por sí mismos carecen de intencionalidad; son bastante carentes de sentido; ni siquiera son manipulaciones de *símbolos*, ya que los símbolos no simbolizan nada. En la jerga lingüística, sólo tienen sintaxis pero no semántica. La intencionalidad que parecen tener las computadoras sólo está en las mentes de quienes las programan y de quienes las usan: los que envían la información de entrada y los que interpretan la información de salida. (SEARLE, 1994: 98)⁸

CONCLUSIONES

1. Luego de efectuar este breve recorrido por algunos de los temas que concentran interés en el ámbito de la filosofía de la mente, es claro que esta joven disciplina filosófica abre sugestivos caminos de circulación para la reflexión en torno a los problemas concernientes a la mente humana y a aquellos aspectos asociados a ella. Los aportes hechos en este campo por aquellos pensadores que hemos visto en esta ocasión constituyen planteamientos bastante representativos de esta problemática e indican el notable nivel de profundidad a que se ha llegado en esta disciplina.
2. Aquel supuesto fundamental de la modernidad en el sentido de que la mente es algo así como el recinto privado de las ideas, instancias radicalmente distintas de lo físico, y en relación con el cual sólo cabría un acceso de tipo privado, ha sido socavado, como hemos visto, por la crítica eliminativista a la psicología popular, sin que ello suponga, claro está, que dicha crítica se encuentre a salvo de cuestionamientos que apuntan a señalar las consecuencias nefastas que podrían producirse a partir de la cristalización de un estado de cosas en que dicho planteamiento adquiriera carta de ciudadanía.

8 En relación con las implicancias de su célebre experimento mental, y las discusiones que en torno a aquéllas se han suscitado, no estará de más considerar los artículos siguientes: Robert Damer (2006), David Cole (documento disponible en la red: < <http://plato.stanford.edu/entries/chinese-room> >; David Anderson y Jack Copeland (2002); Jerome Wakefield (2003).

3. De otro lado, el debate en torno a los contenidos mentales, debate en que se enfrentan el externalismo y el internalismo, pone de relieve un problema agudo: aquél referido al problema de la determinación del significado. Asimismo, a través del esbozo de algunos experimentos mentales concebidos por Putnam y Searle se ha podido apreciar de qué manera se aborda esta cuestión, y en qué términos se sostiene, como ellos lo hacen, que la intencionalidad (aquella virtud intrínsecamente referencial propia de los procesos mentales humanos) posee una irreductible dimensión semántica, que, por ello mismo, supone la articulación de un entramado de relaciones significativas indesligablemente ligadas a la capacidad humana de comprensión de sentidos que el mecanismo simplemente manipulador de símbolos que posee una computadora, por ejemplo, de ninguna manera podría generar.
4. En fin, sobre la base de estos pocos temas, pero que son típicamente representativos del tipo de asuntos tratados en los terrenos de la filosofía de la mente, hemos procurado proporcionar un panorama, ciertamente bastante general, del hervidero de ideas que se está produciendo en aquella disciplina, lo cual nos permite vislumbrar un futuro auspicioso en el sentido de que bien se puede esperar en algún momento que del horizonte de problematización en que ella se inscribe paulatinamente se desprendan mayores luces que hagan posible comprender mejor la enigmática mente humana.

REFERENCIAS BIBLIOGRÁFICAS

- Acero, J. (1995). «Teorías del contenido mental». En M. Broncano (Editor). *La mente humana*. Madrid: Trota, pp. 175-206.
- Anderson, D. y Copeland, J. (2002) «Artificial Life and the Chinese Room Argument». *Artificial Life*, 8, pp. 371-378.
- Boden, M. (1994). «Introducción». En M. Boden (Compiladora). *Filosofía de la Inteligencia Artificial*. México: Fondo de Cultura Económica, pp. 9-32.
- Cole, D. (2004). «The Chinese Room Argument». Recuperado el 4 de mayo de 2008, de <http://plato.stanford.edu/entries/chinese-room>
- Churchland, P. M. (1999). *Materia y conciencia. Introducción contemporánea a la filosofía de la mente*. Barcelona: Gedisa.
- _____ (1981). «Eliminative materialism and the propositional attitudes». *The Journal of Philosophy*. Volume LXXVIII, Nº 2, February, pp. 67-90.

- Churchland, P. S. (2002). *Brain-Wise: Studies in Neurophilosophy*. Cambridge, Massachusetts: The MIT Press.
- _____ (1986). *Neurophilosophy*, Cambridge, Massachusetts: The MIT Press.
- Damper, R. (2006). «The logic of Searle's Chinese Room Argument». *Minds & Machines*. Vol. 16, Issue 2, May, pp. 163-183.
- Fodor, J. (1994) *Psicosemántica. El problema del significado en la filosofía de la mente*. Madrid: Tecnos.
- García, H. (2006). «Searle, John: El misterio de la conciencia» (Reseña). *Lógoi. Revista de Filosofía*. Nº 9, semestre enero-junio, pp. 121-131.
- Gardner, H. (1987). *La nueva ciencia de la mente. Historia de la revolución cognitiva*. Buenos Aires: Paidós.
- González-Castán (1992). «Intencionalidad sin conciencia: Brentano, Searle y las ciencias cognitivas». *Revista de filosofía*. 3ª época, Vol. V, Nº 7, pp. 99-118.
- Martínez, F., P. (2002). *La nueva filosofía de la mente*. Barcelona: Gedisa.
- McCorduck, P. (1991). *Máquinas que piensan. Una incursión personal en la historia y las perspectivas de la inteligencia artificial*. Madrid: Tecnos.
- McCulloch, W. y Pitts, W. (1994). «Un cálculo lógico de las ideas inmanentes en la actividad nerviosa». En M. Boden (Compiladora). *Filosofía de la Inteligencia Artificial*. México: Fondo de Cultura Económica, pp.33-52.
- Putnam. H. (1995). *Representación y realidad. Un balance crítico del funcionalismo*. Barcelona: Gedisa.
- _____ (1988) *Razón, verdad e historia*. Madrid: Tecnos.
- _____ (1975) *Mind, Language and Reality*. Cambridge: Cambridge University Press.
- Rabossi, E. (1995). «La tesis de la identidad mente-cuerpo». En F. Broncano (Editor). *La mente humana*. Madrid: Trota, pp.17-42.
- Rorty, R. (1983). *La filosofía y el espejo de la naturaleza*. Madrid: Ediciones Cátedra.
- Ryle, G. (1967). *El concepto de lo mental*. Buenos Aires: Paidós.
- Sanfélix, V. (2000) «Persona y cosificación (A propósito del sentido práctico de la revolución cognitiva)». En P. Chacón y M. Rodríguez

- (Eds.). *Pensando la mente. Perspectivas en filosofía y psicología*. Madrid: Biblioteca Nueva, pp. 325-356.
- Searle, J. (1994). «Mentes, cerebros y programas». En M. Boden (Compiladora). *Filosofía de la inteligencia artificial*. México: Fondo de Cultura Económica, pp. 82-104.
- _____ (1998). «How to study consciousness scientifically». *Philosophical Transactions: Biological Sciences*. Vol. 353, Nº 1377, 1935-1942.
- Toribio M., J. (2000). «Internalismo, externalismo y ecología». En P. Chacón y M. Rodríguez (Eds.). *Pensando la mente. Perspectivas en filosofía y psicología*. Madrid: Biblioteca Nueva, pp. 231-270.
- _____ (1995). «Eliminativismo y el futuro de la psicología popular». En M. Broncano (ed.). *La mente humana*. Madrid: Trota, pp. 245-272.
- Wakefield, J. (2003). «The Chinese Room Argument reconsidered: Essentialism, indeterminacy and strong AI». *Minds & Machines*, 13, pp. 285-312.
- Wittgenstein, L. (2002). *Investigaciones filosóficas*. Barcelona: Crítica.