

CRÍTICA DE LA RAZÓN COGNITIVA

Jesús Martínez Velasco
Universidad de Valladolid

Resumen. La capacidad de entender, de comprender y de razonar se ha sometido a análisis constantes desde perspectivas a veces contrapuestas, en un intento por elucidar cómo es posible conceptualizar la mente como irreductiblemente psicológica y, a la vez, completamente dependiente de un sistema mecánico causal, como el cerebro. En la actualidad, ha entrado en el escenario una concepción peculiar de la razón, vista desde la perspectiva computacional, que intenta superar tanto el conductismo como el reduccionismo neurofisiológico y se sitúa en el mentalismo clásico, para superarlo, a través de la utilización de las técnicas metódicas propias de la teoría computacional de la mente, cuyo principal objetivo es incidir en el ejercicio de una función con independencia de la sustancia física que le da soporte.

1. INTRODUCCIÓN

La razón, principalmente desde Kant, se ha convertido en tema de estudio y de análisis críticos, observada desde diversos puntos de vista: análisis y crítica de la razón pura, de la razón práctica, de la razón histórica, de la razón instrumental o científico-técnica, de la razón vital, de la razón teórica, de la razón dialéctica, de la razón especulativa, de la razón abstracta, de la razón analítica, de la razón constituyente, de la razón constituida. Hoy, en el marco de la ciencia cognitiva, se somete también a crítica la razón cognitiva.

Desde que Aristóteles definió al hombre como *zoón lógon ejon*, la razón, como facultad definitoria de lo humano, se ha expresado de muchas maneras, pero todas esas diversas formas han tenido como denominador común el intento de alcanzar el conocimiento universal, o universal y necesario, en los distintos campos afectados por la misma.

La razón ha sido el elemento permanente y esencial a lo largo de sus manifestaciones. No sólo la gnoseología racionalista ha situado a la razón en el vértice y en el foco atencional desde el cual iluminar y clarificar todo el proceso de conocimiento. También otros planteamientos gnoseológicos que han insistido en lo que los psicólogos de la percepción llamarían hoy el proceso *bottom-up*, de abajo arriba, de los datos al conocimiento, no han podido prescindir de la razón, aunque, en algunos casos, la hayan reducido a cumplir un papel meramente subordinado, frente al papel principal que se le ha otorgado desde posiciones gnoseológicas racionalistas y, mucho más, idealistas.

La tentación de convertir la razón en un principio hipostático, en una especie de incondicionado absoluto, por ser principio y causa, respecto del cual todo lo demás estaba condicionado y causado, ha hecho perder de vista la dimensión más pegada a la realidad de una razón encarnada y vivida, en la que las experiencias subjetivas adquieren un valor especial difícilmente traducible y menos reductible a meros esquematismos formales y a modelos computacionales o neurales.

Se cree que una visión cientifista del problema de la mente o razón no sería capaz de proporcionar una explicación de aquellos aspectos de la actividad mental que no son fácilmente traducibles a términos operacionales, principalmente de aquellos que se incluyen en el término *qualia*, referentes al carácter fenomenológico de nuestra experiencia, es decir, a la forma en que las cosas aparecen al sujeto consciente. Los *qualia* se han convertido en un problema central en filosofía de la mente, para algunos explicable sólo en términos neurofisiológicos, para otros, aun siendo causalmente dependientes de hechos cerebrales, no se podrían identificar con tales hechos. Para los dualistas, finalmente, ni siquiera son dependientes de los eventos cerebrales, sino que constituyen realidades ontológicamente separables de cualquier realidad física o material. Estas y otras posturas (como la del monismo neutral, la del paralelismo psicofísico, o la del materialismo psiconeural emergentista), no hacen sino confirmar el difícil problema de explicar qué es y, sobre todo, cómo funciona la mente o razón humana.

Desde una perspectiva histórica, hay que considerar meritorio el intento kantiano de superación de una razón pura, buscando en la razón objetiva la síntesis del funcionamiento real de la razón humana. Ni la razón es el poder creador del mundo, ni está sometida de modo constante a los vaivenes de las condiciones empíricas, tanto externas como internas. La razón ha constituido desde siempre el centro de atención de la filosofía, convirtiendo en tema principal de análisis y de estudio de la misma su funcionamiento, sus poderes, sus posibilidades y también sus limitaciones. Y no menos importante ha sido el análisis de su origen, de su estructura y de su configuración. Se han utilizado, a lo largo de la historia, nombres distintos para referirse a esta actividad fundamental del hombre que consiste en su capacidad para comprender, entender, razonar y pensar. Mente, pensamiento, razón y entendimiento han sido y son términos a través de los cuales se ha querido ir precisando el sen-

tido y significado profundo de esa actividad específicamente humana, principalmente en lo que se refiere a uno de los aspectos más subjetivos de las experiencias individuales, la conciencia, como característica esencial de la vida mental.

Tal vez no le falte razón a Searle cuando afirma que el problema mente-cuerpo, mente-cerebro, no sería un problema tan hondo y tan complejo de resolver si no fuera por el problema de la conciencia. Esta da a la actividad de comprender, de razonar, de entender y de pensar una dimensión cuya traducción y reducción a términos neurales y computacionales se ha visto y se ve todavía problemática, y que hoy se traduce en la dialéctica entre razón fenomenológica y razón neuro-computacional.

Por eso, considero que es de actualidad hacer un análisis crítico de la razón cognitiva, porque ésta es la que está cobrando un protagonismo especial en el ámbito de la ciencia cognitiva, que ha tomado como modelo básico del funcionamiento mental la teoría del procesamiento de la información y la metáfora del computador.

2. CRÍTICA DE LA CONCEPCIÓN MECANICISTA DE LA RAZÓN

Cada día se alzan más voces frente a esta concepción mecanicista y objetivista de la mente o razón y, en general, del hombre, porque creen que de la actividad mental o racional y, del hombre, sólo se puede dar una visión fenomenológica, filosófica o sociológica. Algunos verían como peligro o amenaza todo reduccionismo de la racionalidad humana, bien se trate de un reduccionismo conductista, fisiológico o computacional.

Desde una visión humanista del hombre no se acepta como única explicación un proceso causal del mundo natural. Existen sectores de la vida humana, tales como el significado, la intencionalidad, el ejercicio de la capacidad de elección, la intelección, el pensamiento, la propositividad, etc. que se considerarían pertenecientes al ámbito más intrínseco de la vida mental y de los que no se podría dar una explicación desde una perspectiva metodológica positivista.

Este interés por conservar este reducto de positividad de la vida mental no anula, sin embargo, el problema de elucidar "la relación entre una aproximación humanista y mecánica de la mente, de mostrar cómo es posible conceptualizar la mente como irreductiblemente psicológica y, a la vez, completamente dependiente de un sistema mecánico causal, el cerebro"¹. Esta es la tesis que mantiene Searle, aunque, finalmente, se posiciona a favor de un nivel de realidad, el psicológico-subjetivo-consciente, cualitativamente distinto de sus orígenes biológicos e intraducible a términos computacionales.

¹ M. A. BODEN, *Inteligencia artificial y hombre natural*, Madrid, Tecnos, 1984, p. 506.

En el capítulo nueve de *The Rediscovery of the Mind*, Searle critica ciertos aspectos de la exigencia computacional de los estados mentales a los que más adelante nos referiremos con detalle. Por el momento, a modo de síntesis, podemos decir que Searle no acepta que la mente sea un programa de ordenador, pero sí que el cerebro se pueda simular computacionalmente. Acepta, pues, el programa débil de inteligencia artificial y rechaza, por el contrario, el programa fuerte, es decir que la mente sea un programa computacional². No se puede confundir la teoría computacional de la mente con un compromiso con el fisicalismo y con el mecanicismo, porque, como dice Van Gulick³, se puede ser muy crítico con el punto de vista computacional, como es el caso de Searle, y mantener explícitamente un punto de vista materialista de la mente. Bien es cierto que se trata de un punto de vista materialista moderado en el sentido que hemos recordado, referido únicamente a los orígenes neurofisiológicos de dicha actividad, pero no a los aspectos cualitativos que la definen, porque, al igual que rechaza replicar en robots experiencias cualitativas tales como sensaciones, sentimientos y emociones, éstas, cuando surgen como consecuencia de las activaciones de determinadas estructuras y funciones cerebrales, una vez surgidas, son cualitativamente distintas de sus orígenes biológicos.

M. Boden está convencida, sin embargo, de que es posible compatibilizar una visión computacional de la actividad mental con la defensa de una concepción humanista y el miedo a que los modelos mecánicos de la mente nos alienen de nuestra humanidad debe quedar desechado o, por lo menos, menudado, pues, siendo verdad que se pueden aplicar analógicamente a las máquinas términos como subjetividad, significado y propósito, sólo se les puede aplicar de modo secundario⁴.

A pesar de todo, son muchos los que ven en la inteligencia artificial una forma deshumanizante, porque si el hombre tiene que ser modelado como un ordenador, los valores humanos ocuparían un segundo lugar o serían negados. En todo caso, la ciencia natural resultaría inapropiada para comprender la mente. El prejuicio que subyace tiene que ver con la distinción que en la filosofía occidental se ha venido haciendo desde Max Weber y W. Dilthey hasta Habermas entre la comprensión interpretativa (*verstehen*) que ellos ven necesaria para las ciencias sociales y hermenéuticas, y la actitud objetiva, externalizada, propia de la ciencia natural⁵.

Esta dicotomía epistemológica en el terreno global de las ciencias se plasma hoy, en lo que respecta a los problemas de la mente, en la diferencia

² J. R. SEARLE, *The Rediscovery of the Mind*, London, The MIT Press, 1992, pp. 201-202.

³ R. VAN GULICK, "Qualia, functional Equivalence, and Computation", en H. R. OTTO and J. A. TUEDIO, *Perspectives on Mind*, Dordrecht, Reidel, 1988, p. 124.

⁴ M. A BODEN, *Inteligencia artificial y hombre natural*, p. 546.

⁵ M. A BODEN, "Computer Models of Mind: Are they socially Pernicious?", en K. A. MOHYEL-DIN-SAID (ed.), *Modelling the Mind*, Oxford, Clarendon Press, 1990, p. 85.

existente entre una visión de primera o de tercera persona. El predominio cientifista hizo surgir una contra-cultura anticientífica en dominios en los que se suponía que se podían poner en peligro los valores humanos de una especial cualidad. Sin embargo, en psicología, no todos aceptaron la vinculación de esta disciplina al ámbito de las disciplinas en las que sólo era posible proceder metodológicamente desde la comprensión interpretativa, dado que creían que el objeto específico de más interés de la psicología eran las experiencias subjetivas, entendidas como fenómenos de primera persona, únicamente accesibles por introspección y, por tanto, de forma privada. Por el contrario, los paradigmas dominantes en el campo psicológico han sido, por un lado, el conductismo, modelado en torno a las ciencias naturales y, en la actualidad, los modelos neurales y computacionales de la mente, en cuyo desarrollo se han ido perfilando métodos y procedimientos cada vez más precisos para abordar todo el abanico de problemas que plantea la vida psíquica del ser humano. Como “modelos” que son de la vida mental poco se puede decir contra ellos, precisamente por ser modelos, es decir, metáforas o analogías. Todas las ciencias utilizan modelos, algunos incluso para objetivos muy concretos, a veces los modelos son incompatibles unos con otros. Al mismo tiempo, los diferentes modelos sugieren distintas estrategias metodológicas, y cada uno presenta diferentes aspectos del fenómeno en estudio. Esto no sólo ocurre con las ciencias maduras, sino también con las ciencias menos desarrolladas como la psicología⁶.

El peligro puede presentarse cuando queremos convertir “un” modelo en “el” modelo, porque entonces la metáfora o el modelo presupone una teoría y la metodología del modelo se convierte en la metodología de la disciplina. De aquí deriva el miedo que algunos manifiestan cuando se habla hoy del modelo-teoría computacional o del modelo-teoría neural de la mente, porque esos modelos condicionan una manera de entender la realidad humana, próxima para algunos a una visión mecanicista, que impide reconocer y valorar todo el potencial humano. Desde esta premisa, no es extraño que la gente recele, como dice Boden, de teorías psicológicas basadas en el modelo computacional, porque creen que empobrecería nuestra imagen del hombre⁷. Reducir la actividad del pensar a la actividad de una máquina supondría un salto de lo subjetivo a lo objetivo, y un error de primera magnitud. El mismo error que supondría reducir toda explicación a un lenguaje fisicalista, que eliminara toda referencia a la subjetividad, precisamente porque a las máquinas se las consideraría incapaces de acciones con propósitos y únicamente serían procesos ciegos dispuestos por un programador⁸. Ni siquiera admitirían que los procesos psicológicos pudieran ser explicados en términos de procesos neurales de los cerebros humanos porque implicaría reducir la subjetividad y, en general, toda la actividad mental a procesos mecánicos de tipo causal.

⁶ K. V. WILKES, “Modelling the Mind”, en K. A. MOHYELDIN-SAID (ed.), *Modelling the Mind*, p. 63.

⁷ M. A. BODEN, “Computer Models of Mind: Are they Socially Pernicious?”, p. 87.

⁸ M. A. BODEN, *Inteligencia artificial y hombre natural*, p. 564.

El problema de fondo está en negar o admitir la posibilidad de que, desde mecanismos neurales o computacionales, se pueda explicar o no la generación de propiedades mentales, con independencia de las posiciones que finalmente se mantengan, bien sean reduccionistas, fisicalistas, epifenomenalistas, etc. M. Boden, como Searle, no estaría entre ellos. Como Searle, Boden cree que para explicar los fenómenos subjetivos desde el punto de vista causal, hay que acudir al concepto de modelo interno o representación, porque los procesos corporales de nuestros cerebros funcionan como modelos o representaciones del mundo para el individuo de que se trate, de modo que habrá que seguir hablando de significado, propósito y subjetividad para referirse y descubrir a una persona como ser psicológico, aunque dispusiéramos de un conocimiento neurofisiológico completo⁹. No obstante, a diferencia de Searle, Boden cree que el modelo computacional de la mente nos puede ayudar a clarificar conceptos tales como propósito, significado, yo, libertad y elección moral¹⁰, a pesar de las objeciones que habitualmente se presentan a estos modelos a la hora de pretender dar razón desde ellos de determinados fenómenos subjetivos, como los señalados arriba y otros.

Entre otras objeciones, la incapacidad del ordenador, a pesar de su organización aparentemente semejante a la del cerebro, no tanto para reconocer algunos patrones, cuanto para interpretarlos. Igualmente, la imposibilidad de simular en una máquina artificial el conocimiento tácito o los pensamientos indeterminados, característicos del pensamiento humano y que se concreta, de forma expresa, en la falta de tolerancia a la ambigüedad y en la imposibilidad de discriminar entre lo esencial y lo accidental. Al mismo tiempo, se cree que no es posible atribuir emociones a una máquina, como si las emociones carecieran de cualquier componente cognitivo. Por otra parte, hay quienes se apoyan en el teorema de Gödel, según el cual siempre hay límites para lo que cualquier sistema lógico o matemático puede hacer, para llegar a la conclusión de que hay verdades que, por ejemplo, pueden conocer las personas, pero no unas máquinas programadas, porque todo sistema lógico matemático tiene sus límites. Finalmente, apuntaríamos a la siempre recurrida objeción de los límites de la conciencia y, de forma más precisa, de la autoconciencia. Sin entrar por el momento a profundizar en la justificación de estos argumentos que objetan la viabilidad de la vida mental en los ordenadores, conviene adelantar que, aunque algunos de ellos no sean del todo convincentes, sí se puede afirmar que no se pueden simular, por medios computacionales, todos los aspectos del pensamiento humano, y mucho menos hacer una simulación completa, aunque a partir de la arquitectura cognitiva y del funcionamiento de las máquinas actuales se puede concluir, con cierta verosimilitud, que los fenómenos mentales se pueden fundar en una base fisiológica mecánica, a saber: el cerebro¹¹.

⁹ Ibid., pp. 550-51.

¹⁰ Ibid., p. 555.

¹¹ Ibid., p. 572.

No tiene ningún sentido, según Boden, considerar deshumanizante la inteligencia artificial y, sin embargo, haber aceptado como paradigma psicológico el conductismo, cuando éste se proclama como una psicología sin cerebro de corte puramente ambientalista¹². Si a la ciencia natural, con su lenguaje, le resulta difícil expresar cómo la gente es capaz de ver el mundo de un modo diferente, cómo las acciones de una persona difieren de las de otra en función de sus creencias, de sus principios culturales o de sus prioridades personales, por el contrario, pensar la mente como un sistema computacional apoya el punto de vista de que nuestras creencias y valores son psicológicamente cruciales¹³, aunque es evidente que cosas que nosotros hacemos bien, como comprender la lengua materna, reconocer objetos parcialmente ocultos, utilizar el sentido común al abordar un problema, la tecnología de los ordenadores sólo las puede hacer de forma limitada, pero las puede, en alguna medida, simular y, en ese sentido, se puede decir que los modelos computacionales de la mente pueden ser positivamente rehumanizantes. Como añade Boden, la ciencia sólo resultará deshumanizante cuando no haya un espacio para los conceptos mentales, o si no existe un vocabulario para la subjetividad; en cambio, la psicología, la inteligencia artificial y la neurociencia, en la medida en que se orientan sobre las funciones computacionales del cerebro, tienen en cuenta esos conceptos¹⁴.

3. COGNICIÓN Y COMPUTACIÓN

3.1. *Historia*

Gardner, en *La nueva ciencia de la mente*, no duda en afirmar que la filosofía es la más antigua de las ciencias cognitivas y que, desde su vertiente epistemológica, ha proporcionado el programa de trabajo, los temas y las cuestiones en que están trabajando los científicos cognitivistas de orientación empírica.

Los filósofos, desde antiguo, se han formulado muchos interrogantes acerca de fenómenos tales como la representación mental, el pensamiento y su vinculación o no a una realidad mecánica o espiritual, las relaciones entre pensamiento y emoción, etc. Fue Descartes, en la época moderna, el que sentó las bases, de una manera sistemática, de la actual ciencia cognitiva. Se ocupó del pensamiento, de la naturaleza del sistema sensorial, de los procesos cerebrales y de los autómatas como modelos de la naturaleza humana. Pero, dada su posición dualista, nunca imaginó la posibilidad de que la mente se simulase en una máquina, puesto que ni siquiera el autómata-cuerpo era capaz de realizar ninguna operación mental. Atraído como estuvo

¹² Ibid., p. 592.

¹³ M. A. BODEN, "Computer Models of Mind: Are they Socially Pernicious?", 94.

¹⁴ Ibid., p. 96.

siempre por la perspectiva físico-mecánica de Galileo, quiso trasladar esa visión a todo el ámbito de la naturaleza, incluida la naturaleza humana en su dimensión biológica, pero de la que excluía su dimensión mental o psíquica.

En el *Tratado del hombre*, Descartes considera a los animales y al hombre, en su aspecto corporal, meros autómatas mecánicos, pero salvaguarda de esta visión mecanicista aquellas actividades exclusivas y específicas del hombre, irrepetibles e irreproducibles en ningún otro animal o artilugio mecánico, como son todas aquellas que tienen que ver con la *cogitatio*. Tal es la relevancia que Descartes otorgaba a la actividad mental, principalmente en su dimensión consciente, que en *Los Principios de Filosofía* llegó a afirmar que para ser no necesitamos ni de la extensión, ni de la figura, ni de estar en algún lugar, etc, "somos solamente porque pensamos"¹⁵.

El pensamiento es, pues, para Descartes, algo absolutamente inseparable de nosotros, el reducto privilegiado e incorregible de nuestro conocimiento, accesible únicamente mediante la introspección. Descartes separa dos entidades claramente diferenciadas, entre las que la interacción resulta prácticamente imposible, a pesar de los intentos que hace Descartes por explicar cómo los procesos cerebrales podían estar a la base de los procesos psíquicos y cómo la glándula pineal se convierte en la sede de la interacción entre la mente o alma y el cuerpo. No obstante, en las *Meditaciones*, Descartes va a afirmar que el alma o la mente es el único origen de todo pensamiento y creatividad, afirmación que va a retomar Kant cuando, frente a las tesis empiristas, formula lo que se conoce como revolución kantiana, a saber, que en lugar de que el conocimiento se rija por los objetos, son éstos los que se rigen por el conocimiento, aunque con una diferencia substancial con respecto a Descartes, consistente en que el yo kantiano es un ser activo, pero no sobre sí mismo, sino sobre la experiencia. En cualquier caso, el pensamiento es, para Kant, un factor organizador y revelador de todas las posibilidades, anterior a toda experiencia, porque, como había dicho Leibniz, "*nihil est in intellectu quod prius non fuerit in sensu, nisi ipse intellectus*".

A partir de aquí, los planteamientos filosóficos de la filosofía de la mente han ido oscilando desde las tesis defendidas por los miembros del Círculo de Viena, quienes, al aplicar los métodos lógicos al mundo de la experiencia combinaron el espíritu del racionalismo y del empirismo, pasando por G. Ryle, que en 1949 (*The Concept of Mind*) se opuso a la tesis cartesiana de la mente considerándola falsa, porque no tiene existencia separada. Más tarde, R. Rorty critica los intentos de Quine de supeditar la epistemología a la psicología, porque veía en ésta un intento de responder a las cuestiones filosóficas mal concebidas, inclinándose, finalmente, por una explicación neurofisiológica de la experiencia psicológica. Rorty desconfía por completo de la psicología, de la que dice que podía muy bien no haber sido inventada y que, si

¹⁵ R. DESCARTES, *Los Principios de Filosofía*, Madrid, Reus, 1925, p. 26.

desaparece, nadie lamentaría su pérdida. Putnam y Dennett dan un paso más en la formulación de la mente computacional y aunque ninguno de los dos renuncia a sus planteamientos filosóficos, ambos se introducen en la formulación de principios que pueden explicar las manifestaciones cognitivas de los artefactos computacionales, sin que aquellos tengan que reducirse exclusivamente a sus orígenes biológicos, como dice Putnam, o poniendo énfasis, como dice Dennett, en el carácter intencional de los sistemas mentales, carácter que sólo aparece en dichos sistemas a un cierto nivel de complejidad. Por eso, Dennett establecerá una distinción entre intencionalidad intrínseca, propia de los sistemas mentales humanos, y la intencionalidad derivada, característica de los sistemas computacionales.

Pero quien ha situado el tema de la teoría representacional de la mente en un primer plano y dentro del ámbito de la ciencia cognitiva ha sido Fodor, que no ha dudado en reconocer a Descartes el haber admitido la existencia de estados mentales con potencialidad causal. La vinculación al mentalismo racionalista corre paralela con su distanciamiento de posiciones conductistas de origen empirista o neopositivista, como las que representan, por ejemplo, Quine, Ryle y Wittgenstein, y con su alejamiento de cualquier reduccionismo neurofisiológico¹⁶.

3.2. *Fodor y la teoría computacional de la mente*

Fodor piensa que tanto un reduccionismo conductual como fisiológico constituyen una amenaza para la integridad de la teoría psicológica; en ambos casos, el psicólogo saldría perdiendo. Pero, de cualquier manera, cree que se trata de un dilema falso, porque, en contra de Ryle, no piensa que para ser mentalista haya que ser dualista, ya que eso supondría admitir que mentalismo y materialismo se excluyen, cosa que no admite Fodor, porque ni el dualismo ni el conductismo agotan la filosofía de la mente. Si Ryle y Quine no admiten la existencia de cosas tales como creencias, deseos o fenómenos intencionales, Fodor defiende tanto la irreductibilidad como la realidad de estados intencionales, pero, como dice Dennett¹⁷, intentando hacer que estas realidades irreductibles sean aceptables para las ciencias físicas al fundamentarlas en la "sintaxis" de un sistema de representaciones mentales físicamente realizadas.

Fodor trata de ofrecer una justificación fiscalista de la psicología popular como teoría de los estados y procesos intencionales. Estados intencionales que son, de acuerdo con esta teoría, relaciones computacionales con las representaciones mentales, y procesos intencionales que son procesos que implican la manipulación de representaciones mentales por medio de la computa-

¹⁶ J. A. FODOR, *El lenguaje del pensamiento*, Madrid, Alianza, 1975, pp. 24-46.

¹⁷ D. DENNETT, *The Intentional Stance*, Cambridge, The MIT Press, 1987, p. 345.

ción. Estas representaciones mentales pertenecen al lenguaje del pensamiento y a esta teoría se la denomina teoría computacional de la mente (TCM), que es heredera, a su vez, de la vieja teoría representacional de la mente (TRM) y, por consiguiente, heredera también de aquellas explicaciones psicológicas de los eventos mentales entendidos como procesos internos y distinguidos e identificados por sus propiedades funcionales.

Fodor postula una variante materialista de la mente, que admite, sin embargo, la existencia de causas mentales y la interacción mutua de los estados mentales. Las actividades cognitivas se materializan, o se constituyen, en la manipulación de los símbolos o representaciones mentales. En consecuencia, como ya hemos indicado más arriba, se opone a cualquier reduccionismo, a pesar de que los estados mentales se pueden materializar en muchos sistemas. El ideal fisicalista de la ciencia, que ha querido impregnar de lleno el campo de la psicología, no puede, según Fodor, hacernos caer en una visión reductivista de tipo biológico. Según Fodor, la defensa del materialismo o de la unidad de la ciencia no tiene que depender de que para cada estado psicológico deba existir un único estado neural.

Así pues, Fodor considera que el enfoque tradicional del materialismo como el de la unidad de la ciencia están necesitados de una liberalización. Lo que ha pretendido Fodor es dar una explicación de los eventos mentales como procesos internos, distinguidos e identificados por sus propiedades funcionales. Para Fodor, el vínculo entre la mente y el ordenador es más estrecho que el existente entre la mente y el cerebro. Por eso, se manifiesta a favor de la metáfora del ordenador aplicada a los procesos cognitivos y opta por una filosofía de la mente funcionalista, según la cual sistemas tan distintos como los seres humanos, las máquinas de calcular y los espíritus desencarnados tendrían todos estados mentales.

Fodor acepta las tesis fundamentales del funcionalismo de Putnam, según las cuales, la psicología de un sistema depende no de la materia de la que está hecha (células vivas, energía espiritual o metal), sino de cómo funciona y, sobre todo, de su soporte lógico. Las actividades cognitivas se constituyen, o se materializan, en la manipulación de símbolos o representaciones mentales. La mente funciona como un ordenador, pero, a juicio de Fodor, procesando representaciones simbólicas, de acuerdo con las reglas de la lógica. Por eso, el funcionalismo no prevé la eliminación de los conceptos mentalísticos en la explicación de los fenómenos psicológicos y resulta compatible con la solución materialista de las relaciones mente/cuerpo¹⁸. Fodor atribuye a los organismos un sistema representacional, lo que obliga a caracterizar a este sistema como mentalista. Estos símbolos o representaciones no tienen por qué guardar ninguna relación configuracional o de semejanza con las entidades que denotan. Se trata de entidades abstractas, aunque sí poseen contenido

¹⁸ J. A. FODOR, "The Mind-body Problem", en R. WARNER and T. SZUBKA, *The Mind-body Problem*, Oxford, Blackwell, 1994, p. 31.

intencional, porque los símbolos, como los pensamientos, tratan sobre cosas. Esta equivalencia entre lo simbólico y lo mental es la que ha permitido a Fodor mostrar su entusiasmo por la metáfora computacional en el mundo de los procesos cognitivos, ya que ha sido esta metáfora la que ha hecho posible asociar las propiedades semánticas de los estados mentales con las de los símbolos mentales, porque, a su juicio, resulta natural pensar que el ordenador es un mecanismo que manipula símbolos, porque no existe computación sin representación. Los procesos cognitivos son procesos computacionales. Computar es calcular, es establecer relaciones entre símbolos, porque existe un lenguaje del pensamiento que es realizable de forma muy diversa a nivel físico.

Sin embargo, para Fodor, este lenguaje del pensamiento no es un lenguaje natural, sino un pre-requisito para aprender un lenguaje natural. Fodor acepta la existencia de un idioma mental, de un aparato conceptual con que todo individuo viene equipado, capaz de autorizar experiencias e informaciones particulares¹⁹. Siendo verdad que no tenemos conocimiento innato del vocabulario de cualquier lenguaje natural, sí debemos, según Fodor, tener a nuestra disposición un lenguaje no-natural para aprender el primer lenguaje natural que adquirimos²⁰. Hay una revitalización del racionalismo, aunque lo que interesa más a Fodor es una quiebra del empirismo. Se produce una preponderancia del papel del sujeto en el conocimiento que sólo se podrá adquirir en virtud de una estructuración cognitiva previa. Fodor piensa que aunque el pensamiento implica el desarrollo de un sistema lingüístico, es anterior e independiente del lenguaje natural. La equivalencia funcional que establece Fodor entre la mente y el ordenador se basa en que ambos son mecanismos capaces de manipular símbolos mediante el establecimiento de cálculos o computaciones. Pero falta un tercer elemento relacionado con ellos, el cerebro. Así, de acuerdo con Fodor, la mente es un ordenador alojado en el cerebro. Ya Turing intentó poner de relieve la independencia del pensar con su soporte físico. La máquina universal de Turing muestra que hombre y máquinas manipulan símbolos y que el soporte físico no es definitorio para la realización de esta actividad, aunque sí habría que señalar una diferencia fundamental. Si se puede aceptar que los procesos intencionales son procesos computacionales, es decir, procesos que implican la manipulación reglada de oraciones del lenguaje del pensamiento por medio de la computación, hay que añadir, no obstante, que los ordenadores son artefactos mecánicos que no son capaces de comprender el significado de los símbolos que manipulan. Son sensibles únicamente a las propiedades sintácticas de los símbolos. En palabras de Dennett, los ordenadores serían máquinas sintácticas que conllevan (conducen) máquinas

¹⁹ H. GARDNER, *La nueva ciencia de la mente. Historia de la revolución cognitiva*, Buenos Aires, Paidós, 1987, p. 100.

²⁰ Véase J.A. FODOR, *El lenguaje del pensamiento*, cap. 2.

semánticas y, a juicio de Fodor, el cerebro humano sería de la misma naturaleza. Como señala Jackendoff, la teoría computacional de la mente (TCM) se desarrolló a partir de la concepción del cerebro como un sistema de procesamiento de la información, análogo a un ordenador²¹.

El funcionalismo tiene como principal objetivo incidir en el ejercicio de la función con independencia de la sustancia física del cerebro, puesto que lo que se considera principal en el estudio de la mente es precisamente la función más que la estructura física. Resulta importante la información procesada por el cerebro y los procesos computacionales que el cerebro ejecuta sobre esa información, con independencia de cuestiones relativas a su implementación neurológica²².

No se puede confundir la teoría computacional de la mente con un compromiso con el fisicalismo y con el materialismo. Algunos críticos del punto de vista computacional, como Searle, mantienen, sin embargo, de forma explícita, un punto de vista materialista sobre la mente. Lo que importa a los computacionalistas no es tanto formar un sistema con una organización física necesaria, sino que los rasgos más relevantes de esta organización sean rasgos computacionales formalmente especificables²³.

El problema que se plantea es si, desde esa organización sintáctico-formal, se pueden explicar o no las experiencias cualitativas internas, porque, por muy sofisticada que sea la organización funcional de un ordenador, no parece, en principio, que posea esas experiencias internas. Si nos remontamos a Husserl observamos que, a pesar de su formación lógica y matemática, que le familiarizó con una teoría formal o sintáctica, él defiende que las leyes de la lógica se aplican a la descripción fenomenológica de la experiencia en una dirección que no es puramente sintáctica²⁴. Para Husserl, el pensamiento y la experiencia no son puramente formales, lo cual no debería impedir una investigación en inteligencia artificial, pero que si es artificial es porque le faltaría algo de lo que lo define como verdaderamente mental, como sería el carácter subjetivo de las experiencias internas, junto con los aspectos semánticos que definen toda actividad propiamente mental.

Apoyándose en esta distinción entre sintaxis y semántica, los opuestos al modelo computacional de la mente niegan cualquier posibilidad de que se pueda simular realmente en un ordenador el funcionamiento verdaderamente mental de la cognición humana. P. S. Churchland y T. I. Sejnowski²⁵ se

²¹ R. JACKENDOFF, *Consciousness and the Computational Mind*, Cambridge, The MIT Press, 1992, p. 15.

²² Ibid.

²³ R. VAN GULIK, "Qualia, Functional Equivalence, and Computation", p. 124.

²⁴ E. HUSSERL, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*, Erster Buch (Husserliana V), Ed. by Marly Biemal, The Hague, Nijhoff, 1913, §§ 59, 134.

²⁵ P. S. CHURCHLAND and T. I. SEJNOWSKI, "Neural Representation and Neural Computation", en W. G. LYCAN (ed.), *Mind and Cognition*, Cambridge, Basil Blackwell, 1990, pp. 228-33.

hacen eco de estos argumentos para desechar cualquier planteamiento reduccionista en la explicación de la cognición humana y los retrotraen a las relaciones mente/cerebro. Es de sobra conocida la posición de los Churchland al respecto. Su visión naturalista de la mente no plantea ninguna duda sobre su defensa del reduccionismo y en contra de la denominada *folk psychology*, dejando claro que reducción significa, para él, que fenómenos descritos por una teoría se pueden explicar en términos de fenómenos descritos por otra teoría más básica, pero no significa que los fenómenos reducidos desaparezcan ni que sean desacreditados. El argumento principal de los que defienden una posición antirreductiva se funda, sobre todo, en el carácter intencional de los estados psicológicos, que se caracterizan por su contenido semántico, versan y representan otras cosas, y se relacionan semántica y lógicamente. En cambio, las explicaciones neurobiológicas responden sólo a propiedades causales y no pueden ser sensibles a las relaciones lógicas entre los contenidos de los estados cognitivos. De todo ello concluyen que la neurobiología no puede explicar la cognición y, por tanto, la reducción no es posible.

Abundando más a favor del argumento antirreduccionista de la intencionalidad de los estados psicológicos, se trata de justificarlo diciendo que, por un lado, la cognición implica esencialmente representaciones y computaciones. Aquéllas son representaciones simbólicas y éstas reglas para manipular estas estructuras. Por otro, en la explicación semántica de la cognición son esenciales las creencias, los deseos, los pensamientos, etc. Además, estos estados se caracterizan por sus contenidos, tienen intencionalidad y significado y es en virtud de este significado como se puede decir que juegan el papel que juegan en la cognición. Es precisamente en este argumento basado en la intencionalidad en el que se apoya Searle para refutar la teoría computacional de la mente. Veamos.

3.3. Searle y la crítica al modelo computacional

Searle, desde 1980²⁶, está interesado en poner en tela de juicio la tesis según la cual se considera que los procesos mentales son procesos computacionales acerca de procesos formalmente definidos. Su oposición a la teoría computacional de la mente se basa principalmente en el hecho de que un programa de ordenador lo único que puede hacer es manipular sus propios símbolos internos, pero no puede comprender cómo estos símbolos están conectados con la realidad exterior. Precisamente porque nuestros cuerpos, con nuestros cerebros, son máquinas capaces de comprender el significado de esos símbolos es por lo que no podemos parecernos a un programa de ordenador, para el que los símbolos no son símbolos para algo, son sólo señales sin significado. La mente humana es capaz de comprender, no porque se parezca a un ordenador, sino porque, a juicio de Searle, somos un tipo de

²⁶ J. R. SEARLE, "Minds, Brains and Programs", en *The Behavioral and Brain Sciences* 3 (1980) 422.

organismo, poseedor de una cierta estructura biológica, capaz de producir causalmente percepción, comprensión, aprendizaje, acciones, y cualquier otro fenómeno de carácter intencional, entendiéndose por tal “el rasgo de ciertos estados por el que son dirigidos a, o trata sobre objetos y estados de asuntos sobre el mundo”. En este sentido, los estados de un ordenador no son intencionales en el sentido de Searle.

Para Searle, el modelo puramente formal no será suficiente por sí mismo para la intencionalidad porque las propiedades formales no son constitutivas por sí mismas de intencionalidad y carecen de poderes causales, excepto el poder para producir el estadio siguiente del formalismo cuando la máquina está funcionando²⁷. Searle considera que “ningún programa, por sí mismo, es suficiente para la intencionalidad”²⁸. “Porque el programa es puramente formal o sintáctico y porque las mentes tienen contenidos mentales o semánticos, cualquier intento de producir una mente puramente con programas de ordenador abandona los rasgos esenciales de la mente”²⁹.

Si el criterio en el que se basa Searle para atribuir a los estados mentales intencionalidad es que tratan sobre..., que a través de ellos experimentamos las cosas en el mundo, que pensamos sobre..., etc., esto es lo que, a su juicio, falta en los ordenadores, porque resulta inimaginable pensar cómo estos artefactos experimentarían el mundo y menos tener una experiencia consciente del mismo. En definitiva, el problema más importante, para Searle, es el de la conciencia, que ya explicita cuando, refiriéndose a las relaciones mente/cerebro, afirma que ese problema no sería tal sin el problema de la conciencia. Searle, en efecto, establece una estrecha relación entre intencionalidad y conciencia, hasta el punto de que, para él, sólo un ser que tuviera estados intencionales conscientes, tendría verdaderamente estados intencionales y cualquier estado intencional inconsciente sería, por lo menos, potencialmente consciente. A partir de aquí, Searle deduce que cualquier estudio sobre la intencionalidad que deje fuera el tema de la conciencia estará incompleto, porque existe una conexión conceptual entre conciencia e intencionalidad que conlleva el hecho de que una teoría completa de la intencionalidad exige una explicación de la conciencia³⁰.

A este respecto, Jackendoff³¹ responde a Searle que la conciencia no debería convertirse en la justificación para rechazar la mente computacional, aunque sí para limitar el poder de la analogía del ordenador. Searle no sólo limita el poder de esta analogía, sino que rechaza lo que él llama “el programa fuerte de inteligencia artificial”, es decir, se opone a la idea básica del modelo

²⁷ Ibid., p. 422.

²⁸ Ibid., p. 424.

²⁹ J. R. SEARLE, *The Rediscovery of the Mind*, p. 45.

³⁰ Ibid., p. 132.

³¹ R. JACKENDOFF, o.c., p. 20.

computacional de la mente, que consiste en convertir a la mente en un programa de ordenador, como si la mente fuera al cerebro lo que el *software* es al *hardware*, de manera que, además de los problemas de cómo un cerebro puede tener experiencias o desarrollar un razonamiento, se plantea un tercero, a saber, cuál es la relación entre estados computacionales y experiencia. Para Putnam, el problema mente/cuerpo quedaría resuelto si viéramos el cerebro como un sistema de procesamiento de la información. El funcionalismo proporcionaría un modelo explicativo al problema mente/cuerpo computacional, es decir, estructuras y procesos de la mente computacional serían neuralmente ejemplificables. Lo que queda por resolver es si desde esta concepción computacional de la mente se pueden causar o proyectar elementos del saber consciente. Aquí surgen precisamente las divergencias.

Searle, por ejemplo, se manifiesta a favor de que los procesos cerebrales y mentales se puedan simular en un ordenador digital (“programa débil de inteligencia artificial”). Pero pone en duda la afirmación de los cognitivistas de que el cerebro sea un programa de ordenador. Se muestra preocupado por qué tipo de hecho sobre el cerebro constituiría el ser un ordenador³².

Para Searle, no está clara la posición de aquellos para quienes, aun admitiendo que en la mente existe algo más que operaciones sintácticas propias de un ordenador digital, sin embargo, creen que puede ocurrir que los estados mentales sean, por lo menos, estados computacionales, y los procesos mentales sean procesos computacionales que operan sobre la estructura formal de estos estados mentales. Admitido que sea así, admitido que los procesos cerebrales sean procesos sintácticos, que existan oraciones en la cabeza que estén en forma de lenguaje del pensamiento, hay que añadir que, como toda oración o frase, tiene una estructura sintáctica y otra semántica, y ambas pueden separarse. El tema está en cómo esas oraciones en la cabeza logran su significado³³. En este punto es donde Searle marca de nuevo la diferencia entre el carácter puramente formal de un programa y los estados intencionales, cuya definición no les viene de la forma, sino de su contenido. Una creencia no se define como un cierto estado formal, sino como cierto contenido mental con condiciones de satisfacción. Una misma creencia se puede dar en un número indefinido de expresiones sintácticas diferentes en distintos sistemas lingüísticos³⁴.

Así como los estados mentales son producto de la actividad cerebral, un programa no es en ningún caso un producto del ordenador. Para Searle, es definitivo el que un programa de ordenador se limite a manipular símbolos formales desde un punto de vista sintáctico, pero no sabe qué significan los resultados de sus operaciones. Sólo unas máquinas, los cerebros humanos,

³² J. R. SEARLE, *The Rediscovery of the Mind*, p. 204.

³³ *Ibid.*, p. 201.

³⁴ J. R. SEARLE, “Minds, Brains and Programs”, 423.

capaces de causar estados mentales intencionales son las únicas responsables de la actividad mental consciente. Además, esta capacidad causal del cerebro para producir intencionalidad no puede consistir, a su juicio, en una ejemplificación en un programa de ordenador, porque, como ya hemos recordado, ningún programa, por sí mismo, es suficiente para la intencionalidad. Se podrían diseñar robots capaces de hacer las mismas discriminaciones que los robots conscientes, pero serían totalmente inconscientes. Es decir, podríamos formar un robot totalmente inconsciente, pero que tuviera los poderes y habilidades conductuales absolutamente idénticas con respecto a las de un robot consciente³⁵. Esta equivalencia funcional nunca podría llegar a la identificación entre los procesos cerebrales y los procesos computacionales, porque los ordenadores no serían capaces de utilizar el idioma de la intencionalidad y de la conciencia, porque, como decía Quine³⁶, ese idioma no se puede trasladar al lenguaje de la ciencia física. No obstante, a pesar de sus esfuerzos en favor de la unificación de la ciencia y de creer que no hay cosas tales como creencias, deseos o fenómenos intencionales, considera que los idiomas intencionales son indispensables y será problema nuestro ver cómo podemos dar sentido a su empleo en un idioma que Quine califica de dramático, porque cada elemento del vocabulario de ese idioma requerirá una interpretación dramática.

Ya hicimos alusión a que Fodor adoptó una postura anti-Quine y ha defendido tanto la irreducibilidad como la realidad de los estados intencionales, pero intentando, como dice Dennett³⁷, hacer a estas realidades irreducibles aceptables a las ciencias físicas al fundamentarlas en la “sintaxis” de un sistema de representaciones mentales físicamente realizadas. En un punto está de acuerdo Fodor con Searle, y así se lo reconoce éste, a saber, en el hecho de que ejemplificar el mismo programa del cerebro no es una condición suficiente para tener actitudes proposicionales de carácter intencional. A su juicio, se equivocarían quienes, defendiendo el programa de inteligencia artificial, así lo creen³⁸. Sin embargo, le hace dos reproches a Searle. El primero, por no haber comprendido el punto principal sobre el tratamiento de la intencionalidad en el terreno de las teorías representacionales de la mente, que consiste en ver el papel funcional de estas representaciones, y Searle las ve como herederas de las propiedades semánticas de esas representaciones mentales; el problema de la intencionalidad de las representaciones mentales lo reduce, en suma, al problema que confiere propiedades semánticas a un símbolo. El segundo se basa en lo que, a juicio de Fodor, es una confusión, porque no es lo mismo decir que un ordenador (o cerebro) ejecuta operaciones formales sobre símbolos que decir que ejecuta operaciones sobre símbo-

³⁵ J. R. SEARLE, *The Rediscovery of the Mind*, p. 71.

³⁶ W. V. O. QUINE, *Word and Object*, Cambridge, The MIT Press, 1960, p. 221.

³⁷ D. DENNETT, o.c., p. 345.

³⁸ J. A. SEARLE, “Minds, Brains and Programs”, 431.

los formales. Si existen representaciones mentales, deben ser objetos interpretados, y porque son objetos interpretados, los estados mentales son intencionales³⁹.

3.4. *Reivindicación del modelo computacional: niveles de organización*

Según Pylyshyn⁴⁰, Searle nos quiere hacer creer que los ordenadores, en cuanto manipuladores de símbolos formales, carecen necesariamente de intencionalidad o de la capacidad para comprender y para referir, porque tienen diferentes poderes causales de los nuestros. La tesis de Searle es que la intencionalidad está unida a propiedades materiales específicas (cerebrales) y causada por ellas y nada que no proceda de ese origen material podrá tener intencionalidad, como si estuviera proponiendo que la intencionalidad es una substancia extraída (excretada) del cerebro. Para Pylyshyn, sin embargo, tanto las mentes como los ordenadores disponen, desde el punto de vista clásico, de tres niveles de organización⁴¹:

- a) el nivel semántico o nivel de conocimiento, necesario para explicar por qué ciertas metas o creencias tienden a regir ciertas conductas, y por qué éstas pueden variar, de forma racional, si se añaden nuevas creencias u objetivos, estando éstos conectados de forma significativa y racional. Estos cambios semánticos son considerados como signos de una amplia penetrabilidad cognitiva y como diagnóstico de la conducta del nivel de conocimiento requerido.
- b) en segundo lugar, el nivel simbólico, pues el contenido semántico del conocimiento y de las metas está codificado de forma simbólica.
- c) finalmente, el nivel físico o biológico. Por un lado, el sistema entero tiene que realizarse de forma física, porque la estructura y los principios por los que funciona el objeto físico corresponden al nivel físico o biológico. Por otro lado, necesitamos el nivel biológico para explicar cosas tales como los efectos de las drogas o las lesiones del cerebro sobre la conducta, pero también para explicar cosas como la naturaleza del desarrollo cognitivo, o la maduración o la psicopatología, o quizás algunos cambios que ahora se les llama aprendizaje⁴².

Según Newell⁴³, el nivel simbólico requiere un sustrato y la arquitectura es este sustrato definido en un lenguaje descriptivo apropiado. En los seres

³⁹ Ibid., 431.

⁴⁰ Ibid., 442.

⁴¹ Z. W. PYLYSHYN, "Computing in Cognitive Science", en M. I. POSNER (ed.), *Foundations of Cognitive Science*, Cambridge, The MIT Press, 1989, p. 57 y ss.

⁴² Ibid., p. 62.

⁴³ A. NEWELL, P. S. ROSENBLUM and J. E. SAIRD, "Symbolic Architectures for Cognition", en M. I. POSNER (ed.), o.c., pp. 96-97.

humanos, este nivel se concreta en el de circuito neuronal que, de forma corriente, parece bien descrito como redes altamente paralelas interconectadas de conexiones excitatorias e inhibitorias que procesan un *medium* de señales continuas. Por debajo existen otros niveles de descripción, como las neuronas, las macromoléculas..., etc.

Para Pylyshyn, estos tres niveles definen lo que llama el sistema computacional clásico o la arquitectura cognitiva, según el cual tenemos estructuras simbólicas en nuestras cabezas o, como decía Fodor, disponemos de un lenguaje del pensamiento. Existe, sin embargo, una inmensa variedad de arquitecturas y de sustratos físicos en los que pueden implementarse y no resulta fácil comparar diferentes arquitecturas, pues cada una presenta su propio marco total que puede hacer ver el mundo de formas radicalmente distintas. Pues bien, la ciencia cognitiva necesita determinar la arquitectura que subyace y apoya la cognición humana. Por tanto, la arquitectura, en la ciencia cognitiva, constituye un elemento central en una teoría de la cognición humana. Gracias a ella es posible actuar, como dice Newell⁴⁴, de modo flexible en función del entorno, exhibir una conducta adaptativa, usar símbolos y abstracciones, utilizar un lenguaje, tanto natural como artificial, aprender del entorno y de la experiencia, adquirir nuevas capacidades a través del desarrollo, vivir autónomamente dentro de una comunidad social y exhibir autoconciencia y sentido del yo. Esta arquitectura cognitiva, en lo que respecta a la cognición humana, se ha desarrollado por evolución en una tecnología neural y su naturaleza está definida más en términos de funciones que de estructuras y mecanismos, porque, de hecho, una gran variedad de estructuras y mecanismos han sido capaces de suministrar las funciones centrales, como ha puesto de manifiesto la ciencia del ordenador.

El carácter funcional de la arquitectura cognitiva es especialmente importante cuando nos movemos desde los ordenadores digitales hasta la cognición humana. De hecho, ni en unos ni en la otra conocemos todas las funciones que están siendo ejecutadas. En todo caso, la función central de la arquitectura cognitiva es apoyar un sistema capaz de computación universal y sea cual sea la forma que adopte esa arquitectura (neural, computacional, algorítmica, etc.) todas incluyen una manera de formular una máquina universal, a pesar de que el diseño específico de cada una puede hacer oscurecer las funciones generales que se requieren. Ahora bien, ninguna de esas funciones, ni los símbolos que manejan, cumplen la función de representación del mundo externo, porque los símbolos sólo proporcionan una función de representación interna, siendo la representación del mundo externo función del sistema computacional como un todo⁴⁵.

⁴⁴ Ibid., pp. 99-100.

⁴⁵ Ibid., p. 104.

En el lenguaje del pensamiento existen muchos símbolos simples o palabras y reglas sintácticas para combinar estas palabras en orden a formar oraciones. El significado de una oración dada del lenguaje del pensamiento está determinada por el significado de sus palabras o símbolos constituyentes, dentro de una estructura sintáctica y semántica combinatorias. La posición fisicalista de Fodor le lleva a considerar el lenguaje del pensamiento como físicamente encarnado en el sentido de que sus oraciones son idénticas o están constituidas por estados del cerebro, de modo que, para cualquier estado intencional distinto existirá un estado distinto de nuestro cerebro que codifique la oración relevante del lenguaje del pensamiento, oraciones que, por otro lado, son realizables de forma múltiple a nivel físico.

Para Fodor, determinar cómo el lenguaje del pensamiento está codificado en el cerebro de un individuo implica especificar una función mediante la cual se describe la forma física de cualquier símbolo. Esta función es tal que oraciones simbólicamente similares del lenguaje del pensamiento han formado estados del cerebro físicamente similares. Lo que no es posible, como señala Lycan, es que todos los fenómenos mentales se puedan localizar funcionalmente a un mismo nivel, o que cualquier estado mental singular deba ser localizado por entero a un nivel⁴⁶, porque existen distintos niveles en la naturaleza y es un caso especial de funcionalismo el hecho de localizar todos los estados mentales en el nivel más bajo de abstracción, el neuroanatómico, tal como hace la teoría de la identidad, aunque no deja de ser un nivel más de la continuidad de la naturaleza⁴⁷.

Volviendo a Putnam, recordaríamos cómo, para él, el cerebro, visto como un procesador de la información, resuelve el problema mente/cuerpo. Según Putnam, el funcionalismo es capaz de enfocar el problema mente/cuerpo computacional. Así, como recuerda Jackendoff, una solución del problema computacional mente/cuerpo nos diría cómo las estructuras y procesos de información de la mente computacional son neuralmente ejemplificados⁴⁸. La mente computacional se podría, a su juicio, mirar como una especificación abstracta de la organización funcional en el sistema nervioso, porque la mente computacional sería otra manera de describir el cerebro⁴⁹.

Ahora bien, para poder distinguir la mente computacional y la conciencia, Jackendoff es partidario de pasar de la teoría I, según la cual “los elementos del saber consciente *constan de* información y de procesos de la mente computacional que son (1) activos y (2) tienen otras propiedades privilegiadas”, a la teoría II, según la cual, “los elementos del saber consciente están *causados por/apoyados*

⁴⁶ W. G. LYCAN, “The Continuity of Levels of Nature”, en W. G. LYCAN (ed.), o.c., p. 87.

⁴⁷ Ibid., p. 88.

⁴⁸ R. JACKENDOFF, o.c., p. 21.

⁴⁹ Ibid., p. 21.

por/ proyectados desde la información y los procesos de la mente computacional que (1) son activos y (2) tienen otras propiedades privilegiadas”⁵⁰.

La única diferencia entre ambas teorías consiste en sustituir “*consta de*” por “*estar causado/apoyado/proyectado desde...*”. Lo que se puede poner de manifiesto con esta teoría II es, por un lado, la suficiencia computacional, porque afirma que la organización de la conciencia está determinada por la mente computacional que la apoya, y allí donde aparece una diferencia en la conciencia de algo, esta diferencia se corresponde con una diferencia en el estado computacional (ver rojo procede de un estado computacional diferente del de ver azul). En consecuencia, Jackendoff cree que el mundo que nosotros vemos es el que el cerebro/mente computacional es capaz de representar. A su juicio, la teoría computacional debe ser suficientemente expresiva como para hacer el mundo de la conciencia posible, aunque resultaría necesario introducir una condición, en el sentido de que no cualquier distinción computacional causa/apoya/proyecta una distinción fenomenológica, y esta condicionalidad es la que hace posible la computación inconsciente. Por otro lado, se pone de manifiesto la no eficiencia de la conciencia, en el sentido de que la conciencia de cualquier entidad no puede por sí misma tener ningún efecto en la mente computacional. Según esta hipótesis, no existen conexiones causales desde la mente fenomenológica a la mente computacional. La tesis interaccionista queda descartada. Los efectos computacionales sólo se podrían explicar en términos computacionales, pero no al revés⁵¹.

⁵⁰ Ibid., p. 23.

⁵¹ Ibid., pp. 24-25.