

MANUEL DE PINEDO

DEL REPRESENTACIONALISMO  
A LA ARMONÍA PREESTABLECIDA:  
UNA INTRODUCCIÓN CRÍTICA A LA  
FILOSOFÍA DE JERRY FODOR\*

*Resumen.* En este trabajo, tras reseñar una variedad de soluciones al problema mente–cuerpo, analizo con detalle el proyecto de Jerry Fodor de responder a tal problema así como de dar cuenta del fenómeno de la intencionalidad. Señalo que Fodor, al intentar defender simultáneamente la necesidad de estados intencionales para explicar la causación de la conducta y la ineficacia causal de las propiedades intencionales de tales estados, se ve obligado a abrazar una versión de la teoría de la armonía preestablecida entre mente y mundo. Para evitar este resultado, Fodor estaría abocado a abandonar la idea de que los estados intencionales tienen contenido o a reconocer que la causación mental es inevitablemente holista y que, por tanto, no cabe ofrecer una teoría según la cual los estados mentales están determinados por estados cerebrales del sujeto.

*Palabras clave.* Intencionalidad, sobreviniencia, contenido.

FROM REPRESENTATIONALISM  
TO PREESTABLISHED HARMONY:  
A CRITICAL INTRODUCTION TO  
THE PHILOSOPHY OF JERRY FODOR

*Abstract.* In this paper, after reviewing a variety of solutions to the mind–body issue, I analyze in some detail Jerry Fodor’s attempt at answering such a problem and at accounting for the phenomenon of intentionality. I argue that Fodor, by simultaneously defending the need to appeal to in-

tentional states in the explanation of the causation of behaviour and the causally inert character of intentional properties is forced to embrace the idea of preestablished harmony between mind and world. To avoid this result Fodor must either abandon the idea that intentional states have content or recognize that mental causation is inevitably holistic and that, hence, there is no room for a theory according to which mental states are determined by the physical states of the subject.

*Key words:* Intentionality, supervenience, content.

### 1. *Introducción*

¿Cómo es posible el significado y la intencionalidad? ¿Cómo puede haber estados mentales con contenido? Y, ¿cómo pueden jugar un papel en el pensamiento y la acción? Cuando decimos que Murali se ha comprado un lujoso aparato de mini-disk porque quiere escuchar música y porque cree que los aparatos digitales ofrecen la mejor calidad, parece que estamos dando la mejor explicación posible para su acción, y esta explicación apela a estados mentales (deseos, creencias) identificados por su contenido (“Murali oye música”, “los aparatos digitales ofrecen la mejor calidad”). ¿Tienen estos estados poderes causales? Y si es así, ¿funcionan estos poderes causales al margen de las cadenas causales físicas?

En este trabajo estudiaré una respuesta naturalista particularmente influyente a estas preguntas, la hipótesis de Fodor del ‘Lenguaje del Pensamiento’. El proyecto de Fodor es un intento de defender simultáneamente un papel para los estados intencionales en la explicación del pensamiento y la acción, y la idea de que esta explicación es tan legaliforme como las explicaciones de las otras ‘ciencias especiales’ (biología, geología, química, etc.). Fodor consigue esto manteniendo que los estados intencionales pueden ser causa de la conducta en virtud de ser sus características semánticas un reflejo de características sintácticas de representaciones mentales internas, y que estas características sintácticas instancian leyes computacionales. Para él la posibilidad del significado se reduce a la posibilidad de responder a las siguientes dos preguntas: ¿en virtud de qué mecanismos no intencionales puede una representación mental tener contenido (y tener el contenido que tiene)? Y, ¿en

virtud de qué mecanismos no intencionales pueden una representación mental jugar un papel causal? Como veremos, la respuesta de Fodor a la primera pregunta es una teoría causal acerca de cómo las propiedades representacionales covarían con propiedades de aquello que representan. Su respuesta a la segunda pregunta es que las representaciones mentales, que de hecho tienen contenido, son entidades con todo el derecho para aparecer en una teoría científica de la mente porque pueden ser analizadas sintácticamente, y es en virtud de sus propiedades sintácticas por lo que son causalmente eficaces. Fodor también mantiene que las leyes psicológicas (es decir, las leyes que explican el carácter computacional de la interacción entre las características sintácticas de las representaciones mentales) son irreductibles a leyes físicas, aún cuando los ejemplares mentales no son menos físicos que los protones o los planetas.

Mi estrategia será la siguiente: en la sección 2 discutiré una serie de propuestas fisicistas acerca del estatuto ontológico de los estados mentales que finalmente conducen a lo que ha venido en llamarse 'funcionalismo'. En la sección 3 analizaré el proyecto fodoriano, una teoría funcionalista, con cierto detalle. La sección 4 plantea la primera objeción a Fodor: si nos comprometemos con la idea de que los estados mentales son estados *internos* del organismo, necesitamos responder al reto externista que pone precisamente en cuestión este carácter interno de los estados mentales o al menos de su contenido. Estudiaré dos respuestas que Fodor ha intentado dar a este reto, una en esa sección y otra en la 5, y mantendré que ambas involucran o bien abandonar la idea de que las representaciones mentales tienen contenido o bien caer en el 'holismo semántico'. De acuerdo con el holismo semántico u holismo del significado (*meaning holism*), las condiciones de identidad de los estados intencionales dependen de algo más que meras relaciones causales de uno en uno entre estos estados y estados de cosas del mundo; más bien, tales condiciones de identidad dependen de toda la red de estados intencionales. El único recurso que le cabe a Fodor para escapar de esta objeción y de evitar el escepticismo es abrazar una versión de la

idea de que existe una armonía preestablecida entre nuestro pensamiento y el mundo: esto es lo que argumentaré al final de este trabajo.

## 2. *Variedades del fisicismo*

### 2.1. *Del conductismo al funcionalismo<sup>1</sup>*

Existen dos tipos de problemas con los que la filosofía de la mente tiene que enfrentarse: el problema mente–cuerpo y el problema de la intencionalidad. El primero es un problema ontológico y se ocupa de la pregunta “¿cómo puede haber mentes *en* el mundo?”. El segundo es un problema epistemológico (aunque esto mismo es tema de debate: la posición que se discute en este trabajo lo considera puramente semántico) y se ocupa de la pregunta “¿cómo pueden las mentes ser *acerca* del mundo?”<sup>2</sup> El impacto del fisicismo se ha sentido en las respuestas a ambas preguntas, pero de manera más radical en la discusión de la primera. Se entiende aquí como fisicismo el compromiso con la prioridad de la visión ofrecida por las ciencias naturales (que normalmente incluyen, además de la física, la biología, la bioquímica, la neurofisiología, etc.). A grandes rasgos, una posición es fisicista con respecto al problema mente–cuerpo si mantiene que las propiedades mentales dependen constitutivamente, de una forma u otra, de las propiedades físicas (o químicas, o bioquímicas...).

Sin embargo, hay muchas formas de articular esta dependencia. Uno puede defender que tipos de estados mentales son idénticos a tipos de estados físicos, esto es, que las propiedades o características mentales son idénticas a características

<sup>1</sup> Esta sección se ha beneficiado enormemente de discusiones con Tom Beament, Murali Ramachandran y Hilan Bensusan.

<sup>2</sup> La formulación de estas preguntas pierde algo del mordiente de las elegantes versiones inglesas que le debo a Tom Beament: “how can minds be *in* the world?” y “how can minds be *about* the world?”. Por otro lado, se ha sugerido recientemente que el carácter de este problema no es tanto epistemológico como transcendental: no hace referencia a la justificación del conocimiento sino a sus condiciones de posibilidad (Cf., por ejemplo, McCulloch, G., “Phenomenological Externalism” en N. Smith (ed.) *Reading McDowell*, London, Routledge, 2002, pp. 123–139, y también, Wright, C., “Human Nature?” en *Reading McDowell*, London, Routledge, 2002, pp. 14–73).

físicas. Una posición más débil diría que no puede haber cambios en los tipos mentales sin cambios correspondientes en los tipos físicos. Alternativamente, se puede argumentar que, no habiendo correspondencias entre tipos de estados mentales y tipos de estados físicos, cada ejemplar mental es un ejemplar físico. La primera posición se ha llamado ‘teoría de la identidad de tipos’, la segunda mantiene la ‘sobreviniencia’ de lo mental en lo físico y la última es una ‘teoría de la identidad de ejemplares’. A continuación ofreceré un mapa general de las posibilidades y argumentaré contra las más fuertes. A continuación discutiré la idea de sobreviniencia. No ofreceré argumentos críticos contra la idea de identidad de ejemplares, aunque creo que también los hay, y muy poderosos.<sup>3</sup> Señalaré que el abandono de la concepción de lo físico como suficiente para lo mental no involucra rechazar que lo físico puede jugar un papel como condición posibilitadora de lo mental.

La identidad de tipos puede verse como el intento inicial de solucionar el problema mente–cuerpo y puede encontrarse en el conductismo y algunas formas de materialismo reductivo (en este trabajo ‘materialismo’ y ‘físicismo’ son intercambiables). De acuerdo con el conductismo las propiedades mentales son disposiciones a la conducta, siendo estas últimas consideradas como propiedades físicas, como la propiedad de *ser soluble*. Por ejemplo, instanciar la propiedad de *tener un dolor* es una abreviatura de *tener la disposición a hacer esto y aquello en tales y cuales circunstancias*, o *tener la creencia de que p* es una abreviatura de *tener la disposición a responder sí a la pregunta p?*, etc. Ryle, el principal impulsor del conductismo lógico, no niega la existencia de fenómenos mentales. Su ataque al dualismo no es fundamentalmente un ataque ontológico, sino conceptual o categorial. El error que ataca es el de inferir a partir de nuestro uso de categorías especiales para explicar la conducta (“inteligente”, “voluntaria”, etc.) la existencia de otro orden de realidad a la que se aplican. La

---

<sup>3</sup> Ver, por ejemplo, Hornsby, J., “Which Physical Events are Mental Events?”, *Proceedings of the Aristotelian Society* 81, (1980/1), pp. 73–92.

duplicación de sustancias o *res* es fuente de problemas y contradicciones entre las cuales el problema de la interacción es sólo el más notorio. El mito del “fantasma de la máquina” (esto es, el mito de que existe un segundo nivel de entidades dentro del cuerpo físico que explican la conducta) se supera mostrando que las características de las operaciones mentales son realmente características de la conducta. Los conceptos mentales se usan para describir el comportamiento y no para hablar de un conjunto misterioso de “entidades espectrales” (que forman la segunda *res* cartesiana).<sup>4</sup>

Este tipo de identificación tiene que enfrentarse a dos tipos de objeciones. Por un lado, “parecería quedar un residuo intratable de conceptos alrededor de las nociones de conciencia, experiencia, sensación e imagería mental para los cuales alguna explicación que apele a procesos internos sería inevitable”.<sup>5</sup> Las experiencias no son buenos candidatos para explicaciones disposicionales (esta preocupación parece acomodarse mejor dentro de la siguiente teoría de la identidad que comentaré, según la cual las sensaciones y demás son procesos cerebrales). No obstante, el principal problema del conductismo es que existen procesos mentales sin conductas ni disposiciones a la conducta que puedan tratarse como su expresión. Por ejemplo, una explicación conductista de un actor perfecto fingiendo tener un dolor fracasaría o necesitaría apelar a otros estados mentales para explicar el fingimiento. No sólo es problemático que el análisis de las disposiciones parece entrañar un número indefinido de condicionales que cubran los casos innumerables que podrían disparar la conducta en cuestión sino que, más decisivamente, algunos de esos condicionales inevitablemente incluirán un llamada a estados psicológicos no caracterizados conductualmente.

Para mantener el espíritu fisicista del conductismo sin caer en estas dificultades, la opción obvia parecía ser reducir

---

<sup>4</sup> Cf. Ryle, G., *The Concept of Mind*, New York, Penguin, 1963, (1949) (Especialmente los capítulos 1 y 2).

<sup>5</sup> Place, U.T., “Is Consciousness a Brain Process”, en C.V. Borst (ed.) *The Mind/Brain Identity Theory*, London, McMillan, 1970, (1956), pp. 42–43 (Todas las traducciones son mías excepto cuando se indique lo contrario).

los predicados mentales a otros neurofisiológicos en lugar de a disposiciones a la conducta. Esto es lo que el segundo tipo de teoría de la identidad, el materialismo reductor, hace.<sup>6</sup> Los tipos mentales (por ejemplo, *querer una taza de café*) se identifican con propiedades fisiológicas (*tener tal y cual patrón de actividad neuronal*). Esta idea mantiene la pureza de espíritu de la navaja de Occam: “Me parece que la ciencia nos está dando cada vez más un punto de vista según el cual es posible ver a los organismos como mecanismos físico-químicos [...] no existen argumentos filosóficos que nos obliguen a ser dualistas”.<sup>7</sup> Desgraciadamente esta forma de identidad de tipos es también sujeto de objeciones muy duras. Si, de hecho, dos organismos comparten las mismas propiedades mentales pero no las mismas propiedades físicas, la reducción se hace imposible. Esta objeción se debe a Hilary Putnam y supone el golpe de gracia a la teoría de la identidad y, al mismo tiempo, el nacimiento del funcionalismo. Si identificamos propiedades mentales con propiedades físicas nos vemos forzados a decir que dos organismos que no pueden instanciar el mismo tipo de estado físico tampoco pueden instanciar el mismo tipo de estado mental; esto es, las identidades de tipos no permiten la “realizabilidad múltiple”. “De esta forma, si podemos encontrar un solo predicado psicológico que pueda claramente aplicarse tanto a un mamífero como a un pulpo (por ejemplo ‘hambriento’) pero cuyo ‘correlato’ físico-químico sea diferente en ambos casos, la teoría del estado cerebral colapsa”.<sup>8</sup> Deberíamos esperar “que una forma de escribir las leyes psicológicas independiente de la especie pueda ser encontrada”.<sup>9</sup> La conclusión de Putnam es que, por ejemplo, “el dolor

---

<sup>6</sup> La etiqueta ‘teoría de la identidad de tipos’ normalmente se reserva para esta teoría. Sin embargo, esto es ambiguo ya que tanto el conductismo como el materialismo reductor identifican propiedades mentales y propiedades físicas, aunque la elección de estas segundas sea diferente. ‘Materialismo reductor’ conserva esta ambigüedad.

<sup>7</sup> Smart, 1959, pp. 53–54.

<sup>8</sup> Putnam, H., “The Nature of Mental States”, en *Mind, Language and Reality. Philosophical Papers, vol.2*, Cambridge, Cambridge University Press, 1975, (1967), pp. 436–437.

<sup>9</sup> *Ibid.*, p. 437.

no es un estado cerebral, [...] sino un estado de otro *tipo* totalmente. Propongo la hipótesis de que el dolor, o el estado de tener un dolor, es un estado funcional de todo el organismo.<sup>10</sup>

El criterio de Putnam para la identidad de estados mentales es el isomorfismo funcional: dos sistemas son funcionalmente isomórficos si existe una correspondencia entre estados de uno y estados del otro que preserva las relaciones funcionales. La novedad de permitir la existencia de estados mentales entre el *input* sensorial y el *output* conductual que no son reductibles a propensidades conductuales o a propiedades del sistema nervioso permite que estos estados jueguen un papel en la causación de la conducta gracias a su red de relaciones con otros estados mentales. Para el funcionalismo el conjunto de relaciones funcionales (causales) de los estados mentales da su criterio de identidad: dos ejemplares de estados mentales pertenecen al mismo tipo si tienen las mismas relaciones funcionales con otros estados mentales, *inputs* y *outputs*<sup>11</sup> (cf. Putnam 1973b). El modelo seguido por Putnam es el de la teoría de la computación, el modelo de la máquina de Turing: “la misma Máquina de Turing [...] puede instanciarse en un número potencialmente infinito de formas”.<sup>12</sup> El modelo mecánico usado por el funcionalismo no sólo explica la realizabilidad múltiple de los estados psicológicos sino también la identidad de ejemplares entre éstos y los estados físicos.<sup>13</sup>

---

<sup>10</sup> *Ibid.* p. 433.

<sup>11</sup> Cf. Putnam, “Philosophy and Our Mental Life”, en *Mind, Language and Reality. Philosophical Paper, vol.2*, Cambridge, Cambridge University Press, 1975, (1973), pp. 291–303.

<sup>12</sup> Putnam, “The Mental Life of Some Machines”, en *Mind, Language and Reality. Philosophical Paper, vol.2*, Cambridge, Cambridge University Press, 1975, (1967), p. 415.

<sup>13</sup> Cf. Putnam, “Minds and Machines”, en *Mind, Language and Reality. Philosophical Paper, vol.2*, Cambridge, Cambridge University Press, 1975, (1960), pp. 362–385.

## 2.2. *Sobreviniencia*

El funcionalismo no es necesariamente una solución monista al problema mente–cuerpo y, *a fortiori*, no necesariamente fisicista. Una caracterización funcional del pensamiento es compatible con la idea de que la mente está hecha de una sustancia distinta a la que constituye el mundo material (una sustancia no extensa pero organizada funcionalmente) y con la posibilidad de mentes descarnadas que son descritas e identificadas en términos de sus similitudes funcionales. Pero, en la práctica, todas las formas de funcionalismo de las que tengo conocimiento (y, en particular, la de Fodor, que es de la que me ocuparé en este trabajo) también abrazan alguna forma de identidad de ejemplares entre estados mentales y físicos y/o alguna versión de la idea de que la mente ‘sobreviene’ a lo físico aún si no pueden ser identificados.

La tesis de la superveniencia es particularmente difícil de analizar. Para empezar, existen tres caracterizaciones de ella. La primera mantiene que eventos con propiedades de un conjunto M sobrevienen en eventos con propiedades de un conjunto F si dos eventos con propiedades–F idénticas deben también ser idénticos con respecto a sus propiedades–M. He aquí una formulación clásica:

Si dos organismos son microfísicamente indistinguibles el uno del otro, entonces compartirán la misma vida mental; y si dos objetos físicos son microfísicamente indistinguibles el uno del otro también compartirán las mismas macropropiedades [...] [las macropropiedades] “sobrevienen en” o “se siguen de” estados y propiedades microfísicas”.<sup>14</sup>

La segunda caracterización se sigue lógicamente de la primera: si fijas lo físico también fijas lo mental (dado que cambios en M implican cambios en F, si no hay cambios en F no puede haber cambios en M). Una formulación clásica,

---

<sup>14</sup> Kim, J. “Causality, Identity and Supervenience in the Mind–Body Problem”, en F.A. French *et al.* (eds.) *Contemporary Perspectives in the Philosophy of Language. Midwest Studies in Philosophy, vol.IV*, Minneapolis, University of Minnesota Press, 1979, pp. 40-41.

también debida a Kim, es: “La tesis de que lo mental sobreviene en lo físico [es] a grandes rasgos la idea de que el carácter mental de una cosa depende totalmente de su naturaleza física [...]”.<sup>15</sup> La referencia a la “determinación” no tiene porque ser problemática siempre y cuando no se lea como derivabilidad. Es una cosa decir que no pueden darse cambios en lo mental sin que se den cambios en lo físico. Pero diríamos una muy distinta si afirmáramos que las propiedades físicas determinan *cuáles* son las propiedades mentales. Esto es lo que ocurre cuando esta segunda caracterización de la superveniencia se expresa de la siguiente forma: para crear el universo un dios sólo necesitaría preocuparse de las propiedades físicas; una vez que éstas están distribuidas, las propiedades mentales (y éticas, y fenomenológicas, etc.) también lo estarán.

Esta última forma de expresar la idea de la superveniencia asume precisamente el tipo de relación que la tesis de superveniencia intenta explicar. Para fijar lo mental un dios necesitaría hacer algo más que fijar lo físico. También necesitaría estipular las conexiones puente entre los predicados físicos y los mentales. Un universo dado con una distribución específica de propiedades físicas sólo podría tener *una* distribución específica de propiedades mentales (de acuerdo con la tesis de superveniencia), pero nuestro dios no podría determinar *cuáles* son, a no ser que previamente hubiera establecido una correlación entre ambos conjuntos. Si quieres construir un universo con varias capas de propiedades tendrás que leer a Kant y Freud y Kandinsky; con leer a Newton no bastará.

Finalmente, la superveniencia puede explicarse con referencia a sólo un evento u organismo u objeto (es decir, sin la comparación que aparece en la primera de las dos citas de Kim) y sin mención de la determinación. Un ejemplo, esta vez de Davidson (nótese que ofrece dos definiciones, la primera es equivalente a la primera de las de Kim, la segunda no): “no puede haber dos eventos iguales en todos los aspectos físicos

---

<sup>15</sup> Kim, “Supervenience”, en S. Guttenplan (ed.) *A Companion to the Philosophy of Mind*, Oxford, Blackwell, 1994, p. 575.

pero distintos en algún aspecto mental, o [...] un objeto no puede cambiar en algún aspecto mental sin cambiar en algún aspecto físico”.<sup>16</sup> Davidson también ofrece una definición no comparativa de la superveniencia en términos de determinación: “las características físicas de un evento (u objeto o estado) *determinan* las características psicológicas”.<sup>17</sup>

Si dejamos al margen la determinación, podemos ver que hay dos formas de entender la afirmación de Davidson: dos objetos físicamente idénticos deben ser mentalmente idénticos; y no puede haber cambios mentales sin cambios físicos. La segunda es más débil dado que permite que los cambios ocurran en cualquier sitio y no necesariamente en las características o propiedades físicas del organismo al que se adscribe mentalidad.

Las posibilidades de finura de análisis son desorbitadas. Para empezar, la superveniencia puede aplicarse dentro de un mundo posible (‘superveniencia débil’) o a través de todos los mundos posibles (‘superveniencia fuerte’). Por otra parte, y es sobre ésta sobre la que diré algo más, puede distinguirse entre ‘superveniencia estrecha’ y ‘amplia’ (llamadas usualmente en la literatura ‘local’ y ‘global’). Todas las citas de Kim y Davidson que he ofrecido son paradigmáticas de la superveniencia estrecha. Dados dos organismos físicamente iguales, cualquier propiedad mental que uno posea también la posee el otro (cuáles sean éstas es otra cuestión, como he señalado más arriba). Para la superveniencia amplia no puede haber cambios mentales sin cambios físicos, pero basta con que haya cualquier cambio físico en el universo para que pueda darse un cambio mental. He aquí algo que parece una formulación alternativa de la superveniencia amplia: dos mundos posibles no pueden ser

---

<sup>16</sup> Davidson, D. “Mental Events”, en *Essays on Actions and Events*, Oxford, Clarendon Press, 1980, (1970), p. 214.

<sup>17</sup> Davidson, “The Material Mind”, en *Essays on Actions and Events*, Oxford, Clarendon Press, 1980, (1973), p. 253. La formulación de Davidson tiene un problema adicional: asume que *el mismo* evento o estado tiene características físicas y psicológicas. No es claro que la aceptación de la tesis de la superveniencia implique la aceptación de la identidad de ejemplares. Cf. Haugeland, J., “Weak Supervenience”, en *Having Thought. Essays in the Metaphysics of Mind*, Cambridge, Mass., Harvard University Press, 1998, (1982), pp. 94 y ss.

físicamente indistinguibles pero mentalmente diferentes.<sup>18</sup> Al igual que en lo referente al paso de determinación a derivabilidad, este paso me parece insatisfactorio: no veo objeción a concebir dos universos físicamente idénticos pero con distintas correlaciones entre propiedades mentales y físicas. No tengo conocimiento de una distinción entre una versión débil y una fuerte del principio de superveniencia global o amplio, pero la última formulación atribuida a Kim abre la posibilidad de esta distinción: uno podría desear mantener una concepción de lo físico como condición necesaria o posibilitadora de lo mental, pero no suficiente (de ahí la etiqueta ‘superveniencia débil global/amplia’). No obstante, será sólo la superveniencia estrecha la que sea objeto de escrutinio en este trabajo ya que es ésta la aceptada por Fodor y a la que opondré los famosos experimentos mentales externistas ofrecidos por Putnam y Burge.

Para terminar esta breve tarea de limpieza ontológica, mencionaré brevemente una última opción fisicista, la identidad de ejemplares. Frente a las tesis consideradas en los párrafos anteriores, la identidad de ejemplares ignora el problema de la relación entre propiedades mentales y físicas y sólo afirma que cada ejemplar de un estado mental es idéntico con un estado físico. En general esta idea va acompañada por una teoría que explica cómo las entidades mentales pueden heredar los poderes causales de las físicas. Este es el caso de Davidson,<sup>19</sup> aunque no me detendré más en él.

### *2.3. De vuelta al funcionalismo*

Como hemos visto, el funcionalismo es en la práctica, aunque no por principio, una teoría fisicista comprometida con la existencia de estados mentales que ofrece una explicación de la realización múltiple de estos estados: organismos con estructuras físicas muy diferentes pueden compartir estados mentales y el mismo organismo puede estar en estados físicos diferentes y, sin embargo, cada uno de ellos puede

---

<sup>18</sup> Kim, “Supervenience”, en S. Guttenplan (ed.) *A Companion to...* cit., p. 578.

<sup>19</sup> Cf. Davidson, “Mental Events”, en *Essays on Actions...* cit., pp. 207–225.

instanciar el mismo estado mental (tipo).<sup>20</sup> Ahora bien, como Putnam mostró en una serie de influyentes artículos en la primera mitad de los setenta,<sup>21</sup> la identidad de estados físicos internos no es condición suficiente para la identidad de estados psicológicos (también mostró la década anterior, inventando así el funcionalismo, que esa identidad tampoco era una condición necesaria, como hemos visto antes). Dos organismos pueden compartir sus estados cerebrales y encontrarse en estados funcionales diferentes, dado que el medio y la comunidad lingüística a la que pertenezcan también determinan su contenido. La segunda parte de este trabajo se ocupa de esta crítica y de sus consecuencias para la respuesta a la segunda de las preguntas mencionadas al principio de esta sección, el problema de la intencionalidad. Pero ahora pasaré a analizar una teoría funcionalista particularmente influyente, la de Fodor.

### 3. *La Teoría Representacional/Computacional de Fodor*

Todo el sistema de Fodor se construye a partir de dos axiomas básicos:

- (1) El realismo intencional, y
- (2) El fisicismo.

Para Fodor, la principal implicación de (1) es un compromiso ontológico con la existencia de estados mentales (creencias, deseos y demás) que tienen la estructura de actitudes proposicionales y forman una red de estados intencionales nomológicamente relacionados. Esto supone la introducción de un tercer conjunto de entidades entre las dos reconocidas por los conductistas, *input* y *output*. “[A] menudo es parte de la expli-

---

<sup>20</sup> Cf. Fodor, J. “The Present Status of the Innateness Controversy”, en *Representations*, Brighton, Harvester Press, 1981, pp. 9–11; y también ver Fodor and Block, N., “What Psychological States Are Not”, en *Representations*, Brighton, Harvester Press, 1981, (1972), pp.79–99.

<sup>21</sup> Cf. Putnam, “Meaning and Reference”, *Journal of Philosophy* 70, (1973), pp. 699–711, y también ver Putnam, “The Meaning of ‘Meaning’”, en *Mind, Language and Reality. Philosophical Paper, vol.2*, Cambridge, Cambridge University Press, 1975, pp. 215–271.

cación de la manera en la que un organismo se comporta mencionar las creencias que tiene sobre el tipo de comportamiento que produce”.<sup>22</sup> Es más, Fodor duda que podamos explicar la conducta sin usar teorías intencionales.<sup>23</sup> Este compromiso con los conceptos intencionales implica un compromiso con los métodos de la psicología de sentido común (*folk psychology*), donde estos conceptos tienen su origen. Mantiene Fodor, contra eliminativistas como Stich o Paul y Patricia Churchland, que la psicología de sentido común es una teoría tan buena que a menudo no somos conscientes de que estamos usándola.<sup>24</sup> Como es sabido, el principal argumento de Fodor a favor de construir una teoría científica de la intencionalidad es la falta de alternativas.<sup>25</sup> No me ocuparé aquí de este asunto ya que mi propósito es mostrar como la cadena inferencial que lleva de (1) al resto de las tesis de Fodor se ve bloqueada y no cómo debería comprenderse (1).

La idea de que los estados mentales son materiales le permite a Fodor defender una tercera y fundamental tesis:

(3a) Los estados mentales tienen poderes causales.

Esta tesis, y sus consecuencias, es la clave para poder utilizar estados intencionales para explicar la conducta. (3a) puede defenderse sin comprometerse con (2). Pero esto con-

<sup>22</sup> Fodor, *The Language of Thought*, New York, Thomas Y. Crowell, 1975, p. 28  
<sup>23</sup> Fodor, *Psychosemantics*, Cambridge, Mass., The MIT Press, 1987, p. xii.

<sup>24</sup> Por ejemplo, en *Psicosemántica*, dice que “[l]a psicología de sentido común funciona tan bien que desaparece”. *Ibid.* p. 3. La idea es que no somos conscientes de usarla para predecir y explicar la conducta porque casi nunca tenemos que cuestionar su aparato teórico. La afirmación podría invertirse y dejaría de servir como argumento intuitivo a favor de las tesis de Fodor: no es porque funciona bien por lo que desaparece, sino que funciona bien porque no aparece. Claro está, Fodor le saca mucho partido al carácter teórico de la psicología folk. Una teoría científica de la mente tendría que hacerse cargo de esta teoría preexistente o bien sustituyéndola (como quisieran los eliminativistas) o usándola como punto de partida. Alternativamente, uno puede negarle el status de teoría a la psicología de sentido común y oponerse al proyecto de una ciencia de la mente. Espero que estas anotaciones dejen claro que mi uso de “psicología” folk no tiene ni las connotaciones negativas que los eliminativistas relacionan con “folk”, ni las connotaciones científicas que Fodor relaciona con “psicología”.

<sup>25</sup> *Ibid.*, pp. 1-16.

duciría a rechazar otra tesis que aparece en el sistema de Fodor y en gran parte de la filosofía moderna, la idea de que los poderes causales han de ser subsumibles por leyes y, por ello, ser en última instancia materiales: “cualquier cosa que tenga poderes causales es ipso facto material”.<sup>26</sup>

(3b) La causalidad debe ser, en última instancia, capturable por leyes.

De hecho, las actitudes proposicionales, las creencias, deseos y otros conceptos de la psicología folk, no serían útiles a la psicología científica si no fueran físicas, si se asume que no hay causalidad excepto la causalidad física. El resto del sistema psico-filosófico de Fodor es un intento de desarrollar la idea de que los estados mentales son causas nomológicas de la conducta, esto es, a desarrollar (3) (en mi discusión trataré (3a) y (3b) como una sola tesis, a la que llamaré (3), ya que él no las considera separables).

El siguiente paso es:

(4) La Teoría Representacional de la Mente: un estado mental es una relación con una representación y un proceso mental es un proceso en el que representaciones son transformadas.

Aquí las consecuencias de (3) empiezan a salir a la luz. He citado antes un pasaje en el que Fodor declara que necesitamos hacer referencia a las creencias del organismo para explicar su conducta. Y, de acuerdo con él, esto sólo puede hacerse si “asumimos que el agente tiene medios para representarse su conducta a sí mismo”.<sup>27</sup> Por tanto, “ciertos tipos de patrones centrales de explicación psicológica presuponen la accesibilidad, para el sistema que actúa, de alguna forma de sistema representacional”.<sup>28</sup> Más exactamente, la teoría representacional de la mente es la conjunción de las dos tesis siguientes:

---

<sup>26</sup> *Ibid.*, p. x.

<sup>27</sup> Fodor, “*The Language of...*”, cit., p. 30.

<sup>28</sup> *Ibid.*, p. 31.

*Tesis 1* (la naturaleza de las actitudes proposicionales):

Para cualquier organismo O, y cualquier actitud A hacia la proposición P, hay una relación ('computacional'/'funcional') R y una representación mental MP tal que

MP significa que P, y

O tiene A syss O tiene R con respecto a MP.

*Tesis 2* (la naturaleza de los procesos mentales):

Los procesos mentales son secuencias causales de instancias (*tokenings*) de representaciones mentales.<sup>29</sup>

La teoría representacional de la mente obtiene todo su poder en combinación con la metáfora computacional, hacia la cual el funcionalismo está tan inclinado:

(5) Teoría Computacional de la Mente: tener una actitud proposicional es "estar en alguna relación computacional con una representación interna".<sup>30</sup>

"[E]l acto ejecutado por el agente es la consecuencia de las computaciones definidas a partir de representaciones de posibles acciones. Sin representaciones no hay computaciones. Sin computaciones no hay modelo".<sup>31</sup> El modelo computacional explica cómo las representaciones (que, obviamente, tienen propiedades semánticas por el hecho de ser representaciones de algo) pueden tener poderes causales. La idea es que tener una actitud proposicional involucra instanciar un símbolo y que "[c]onectas las propiedades causales de un símbolo con sus propiedades semántica *por medio de su sintaxis*"<sup>32</sup> porque "sabemos gracias a la lógica moderna que ciertas relaciones semánticas entre símbolos pueden, por así

<sup>29</sup> Fodor, "*Psychosemantics...*", cit., p. 17.

<sup>30</sup> Fodor, "*The Language of...*", cit., p. 198.

<sup>31</sup> *Ibid.*, p. 31.

<sup>32</sup> Fodor, "*Psychosemantics...*", cit., p. 18.

decirlo, “ser imitadas” por sus relaciones sintácticas”<sup>33</sup>. Y esta es la siguiente tesis de Fodor. Parafraseando a Dennett:

(6) La mente es un motor sintáctico.

Este es un momento esencial en la construcción del sistema fodoriano. Aquí es donde la herencia de Chomsky se hace notar con mayor claridad (pero, ver también las tesis 7, 8, 9 y 12): la tesis de Chomsky de la independencia de la gramática es una asunción de fondo en la filosofía de Fodor. Por ejemplo, reconoce en el prefacio de *The Psychological Explanation*<sup>34</sup> que su trabajo es un intento de hacer explícitas las consecuencias para la psicología de la teoría generativa de la gramática. Chomsky argumenta que no podemos explicar toda la conducta lingüística como resultado de convenciones, al menos no en el caso de la sintaxis (como Searle ha señalado,<sup>35</sup> Chomsky concibe al hombre como animal sintáctico). Mantiene que el aprendizaje de nuestra lengua materna ocurre antes y más rápido que el aprendizaje en otras áreas. De hecho, los datos a los que estamos expuestos en nuestro periodo de aprendizaje son inadecuados para explicar lo rápido y prematuro del mismo: “Bajo circunstancias normales aprendemos palabras por medio de una exposición limitada a su uso”<sup>36</sup>. El proyecto de Chomsky está dirigido por:

“...la posibilidad de que estudiando el lenguaje podamos descubrir principios abstractos que gobiernan su estructura y uso, principios que son universales por necesidad biológica y no un mero accidente histórico, que se derivan de las características mentales de la especie.”<sup>37</sup>

En este sentido técnico, “la teoría lingüística es mentalista, ya que se ocupa de descubrir una realidad mental por

---

<sup>33</sup> *Ibid.*, p. 19. “[L]a estructura sintáctica de un estado mental refleja las relaciones semánticas entre sus objetos intencionales” *Ibid.* p. 138.

<sup>34</sup> Cf. Fodor, *The Psychological Explanation*, New York, Random House, 1968.

<sup>35</sup> Searle, J. “Chomsky’s Revolution in Linguistics”, en G. Harman (ed.) *On Noam Chomsky*, New York, Anchor Books, 1974, (1972), p. 15.

<sup>36</sup> Chomsky, N., *Problems of Knowledge and Freedom*, London, Fontana house, 1971, p. 22.

<sup>37</sup> Chomsky, *Reflections on Language*, London, Fontana, 1976, p. 4.

debajo del hecho de la conducta”.<sup>38</sup> No hay razones para pensar que pueda encontrarse una teoría del aprendizaje que explique la adquisición del lenguaje (y de las estructuras cognitivas en general) a partir de la experiencia. Para Chomsky esto significa una defensa de la ‘hipótesis del innatismo’ y que la siguiente pregunta es “¿existe tal cosa como una teoría del aprendizaje, a la espera de ser descubierta?”<sup>39</sup> Su respuesta es que no; a diferentes dominios cognitivos le corresponden diferentes teorías explicativas: “Puede bien ser el caso, como ha menudo se ha especulado en el pasado, que la facultad humana del lenguaje sea relativamente independiente de otras facultades de la mente y también un atributo únicamente humano”.<sup>40</sup>

La defensa reiterada de la sintacticidad de la mente por parte de Fodor es resultado, no siempre explícito, de la necesidad de separar un campo innato de entidades que puedan explicar el aprendizaje lingüístico. Así, aunque la tesis del innatismo (ver (12)) es una consecuencia metodológica de (6), (6) se postula porque lo único que puede ser innato es la sintaxis, la gramática; el significado y la referencia sólo aparecen cuando el lenguaje es “usado”, esto es, cuando la ‘competencia’ entra en relación causal con el mundo y se convierte en ‘actuación’ (*performance*). Otra razón por la que esta tesis es importante para nosotros es que, como intentaré mostrar, la restricción sintáctica que Fodor impone a la teoría de la mente hace que el significado sea, en última instancia, causalmente inerte (es decir, el contenido de un estado mental es parte de sus condiciones de identidad, pero la mente no es causalmente eficaz en virtud de él).

Fodor aún necesita algo más para llegar a su idea programática de un ‘lenguaje del pensamiento’ (ver (8)):

---

<sup>38</sup> Chomsky, *Aspects of the Theory of Syntax*, Cambridge, Mass., The MIT Press, 1965, p. 4.

<sup>39</sup> Chomsky, *Reflections on Language*, London, Fontana, 1976, p. 14.

<sup>40</sup> Chomsky, *The Logical Structure of Linguistic Theory*, New York, Plenum Press, 1975, p. 10.

(7) Los estados mentales tienen una estructura constitutiva.

No sólo los objetos de los estados intencionales (los contenidos) son complejos, los estados mismos también lo son. Para Fodor no hay ‘mecanismos intencionales’, es decir, el contenido *per se* no determina el rol causal. Podemos hablar de causalidad mental en términos de la constitución sintáctica de los estados mentales, sin referirnos a sus propiedades intencionales.<sup>41</sup> Una vez que se ha probado que los estados intencionales son representacionales y que tienen un papel causal sólo en virtud de su constitución sintáctica, Fodor llega a la que seguramente sea su tesis más original:

(8) Los sistemas representacionales pueden ser comprendidos como lenguajes del pensamiento innatos.

La psicología cognitiva “trata a la conducta como el resultado de la computación y la computación presupone un medio en el que se computa”;<sup>42</sup> “las representaciones presuponen un medio representacional y no hay simbolización sin símbolos. En particular, no hay representación interna sin un lenguaje interno”.<sup>43</sup> El lenguaje del pensamiento no puede ser un lenguaje natural, porque organismos no lingüísticos o pre-lingüísticos también piensan y porque necesitamos, de acuerdo con Fodor, un lenguaje no aprendido para aprender nuestra primera lengua (natural).

Fodor acepta así lo que Wittgenstein ha criticado bajo el nombre de ‘imagen agustiniana del lenguaje’:

Agustín describe el aprendizaje del lenguaje humano como si el niño llegase a un país extraño y no entendiese el lenguaje del país; esto es: como si ya tuviese un lenguaje, sólo que no ese. O también: como si el niño ya pudiera *pensar*, sólo que no todavía hablar. Y “pensar” querría decir aquí algo como: hablar consigo mismo.<sup>44</sup>

<sup>41</sup> Fodor, “*Psychosemantics...*” cit. p. 139.

<sup>42</sup> Fodor, “*The Language of...*” cit., p. 33.

<sup>43</sup> *Ibid.*, p. 55.

<sup>44</sup> Wittgenstein, L., *Investigaciones filosóficas*, México, Crítica, 1986, (1953), p. 32. (Traducción de *Philosophische Untersuchungen* por Alfonso García Suárez y Ulises Moulines).

Fodor ofrece su teoría del lenguaje del pensamiento como una teoría de las actitudes proposicionales. En su artículo “Propositional attitudes”<sup>45</sup> dice que las condiciones a priori que esta última teoría tiene que cumplir son: (i) las actitudes proposicionales deben ser analizadas como relaciones (esto es intuitivamente correcto; la única alternativa es considerarlas “fundidas” a su objeto y esto, según Fodor, no tiene futuro); (ii) los verbos de actitud proposicional son paralelos a verbos de decir; (iii) las actitudes proposicionales son opacas; (iv) los objetos de las actitudes proposicionales tienen forma lógica; y, finalmente, (v) debe encajar con las explicaciones empíricas de procesos mentales. Carnap<sup>46</sup> sugirió construir las actitudes proposicionales como relaciones entre personas y las oraciones (del lenguaje natural) que tuvieran la disposición de proferir. De acuerdo con Fodor, esta propuesta, que desplaza de un lenguaje interno e innato a uno público la explicación de las actitudes proposicionales, no puede funcionar porque: (a) puedes creer algo sin hablar ningún lenguaje; (b) existen creencias inexpresables; y (c) si los ejemplares de oraciones de un lenguaje natural son los objetos de las actitudes proposicionales, ¿cómo puede aprenderse la primera lengua? Todo lo que se necesita para hacer justicia a las restricciones (i)–(v) es que los objetos de las actitudes proposicionales sean oraciones, de forma que la manera de evitar (a)–(c) es considerar que no son oraciones de un lenguaje natural, sino oraciones/fórmulas de un sistema interno de representaciones.

Hay otras cuatro tesis importantes defendidas por Fodor que aparecen o bien como consecuencias de las ya discutidas o como recursos teóricos para evitar las críticas que veremos más adelante:

(9) La modularidad de la mente,

(10) El solipsismo metodológico,

---

<sup>45</sup> Cf. Fodor, “Propositional Attitudes”, en *Representations*, Brighton, Harvester Press, 1981, (1978), pp.177–203.

<sup>46</sup> Cf. Carnap, R., *Meaning and Necessity*, Chicago, University of Chicago Press, 1947.

(11) La semántica denotacional/informacional,

(12) El innatismo de todos los conceptos.

La tesis (9) es una consecuencia de la teoría del aprendizaje lingüístico que se sigue de (8), y tiene una profunda influencia chomskyana. Sin entrar en detalles, la idea general es que la percepción y el lenguaje deben considerarse módulos independientes de otras formas de procesamiento de la información. Estos módulos están especificados genéticamente, son autónomos y están relacionados con sistemas neuronales específicos.<sup>47</sup> Fodor sigue las ideas de Chomsky sobre el aprendizaje lingüístico que, como hemos visto, es específico con respecto al dominio.

La tesis (10) es muy importante para Fodor. Una posible glosa es decir que debemos aproximarnos a la tarea de la psicología sin tener en cuenta nada externo al individuo, es decir, haciendo referencia solamente a la contribución del organismo al contenido de sus estados mentales, al ‘contenido restringido’ (*narrow content*). (10) puede considerarse una consecuencia de (6):

[s]i los procesos mentales son formales, sólo tienen acceso a las propiedades formales de las representaciones del medio que los sentidos proporcionan. Por tanto, no tienen acceso a las propiedades *semánticas* de tales representaciones, incluyendo la propiedad de ser verdadero, de tener referentes o, de hecho, a la propiedad de ser representaciones *del medio*.<sup>48</sup>

La aceptación de la teoría computacional de la mente (o, lo que en este contexto es lo mismo, de la condición de formalismo) conduce al solipsismo metodológico como estrategia de investigación en la ciencia cognitiva. Volveré en secciones posteriores a los enormes problemas que esto acarrea.

La tesis (11) es una de las que han ocupado más tiempo a

---

<sup>47</sup> Para una explicación detallada de este modelo ver Fodor, *The Modularity of Mind*, Cambridge, Mass., The MIT Press, 1983.

<sup>48</sup> Fodor, “Methodological Solipsism Considered as a Research Strategy in Cognitive Science”, en *Representations*, Brighton, Harvester Press, 1981, (1980) p. 314.

Fodor en los últimos años. Surge directamente de (4) y mantiene que una teoría semántica, o una teoría del contenido, sólo es posible por medio de una explicación causal refinada, una teoría emparentada con las ofrecidas inicialmente por Stampe<sup>49</sup> y Dretske<sup>50</sup> y no por medio de una teoría del rol conceptual o una semántica fregeana de dos factores ('sentido' y 'referencia') que conducen al holismo semántico. La motivación fundamental para una semántica denotacional es precisamente evitar esta objeción de holismo semántico. Su búsqueda de este tipo de semántica está íntimamente conectada con su concepción del programa de naturalización de la semántica. La naturalización de la semántica (esto es, la naturalización del vocabulario primitivo, no lógico, del lenguaje del pensamiento en términos no intencionales y no semánticos) es una consecuencia de (2) y se logrará, siempre según Fodor, por medio de relaciones causales entre este vocabulario y las propiedades expresadas por él en un contexto dado. Tendremos tiempo de volver a este punto.

La tesis (12) es una de las primeras consecuencias de (8) sobre las que Fodor llamó la atención. Ha sido considerada por muchos autores<sup>51</sup> como una reducción al absurdo de la teoría del lenguaje del pensamiento: “[s]i los mecanismos de aprendizaje conceptual son la proyección y la confirmación de hipótesis (y qué otra cosa podrían ser), entonces existe un sentido en el que no puede darse tal cosa como el aprendizaje de un nuevo concepto”.<sup>52</sup>

Fodor considera que la “demostración” por parte de Chomsky de que existen claros indicios del carácter innato de una ‘teoría lingüística universal’ es el único resultado importante en la ciencia cognitiva. Por ejemplo, en “The present

---

<sup>49</sup> Cf. Stampe, D., “Towards a Causal Theory of Linguistic Representation”, *Midwest Studies in Philosophy* 2, (1977), pp. 42–63.

<sup>50</sup> Cf. Dretske, F., *Knowledge and the Flow of Information*, Cambridge, Mass., The MIT Press, 1981.

<sup>51</sup> Ver por ejemplo Putnam, *Representation and Reality*, Cambridge, Mass., The MIT Press, 1988.

<sup>52</sup> Fodor, “*The Language of...*” cit., p. 95.

status of the innateness controversy”,<sup>53</sup> dice que sólo una teoría se ha propuesto para el aprendizaje conceptual y que incluye la fijación inductiva de ciertas creencias; los mecanismos para esta fijación son la formulación y confirmación de hipótesis acerca de los conceptos que se aprenden. Pero, y éste es el principal argumento de Fodor a este respecto, una teoría del aprendizaje conceptual no puede ser una teoría de la obtención o adquisición de conceptos, porque los conceptos deben ser accesibles para la formación de las hipótesis. Los conceptos son “disparados” por la experiencia, pero no aprendidos por medio de la experiencia.

[U]na teoría de la plasticidad conceptual de los organismos debe ser una teoría acerca de cómo el medio selecciona entre los conceptos especificados de forma innata. *No es una teoría acerca de cómo adquieres los conceptos, sino una teoría acerca de cómo el medio determina qué partes de los mecanismos conceptuales en principio accesibles para ti de hecho se explotan.*<sup>54</sup>

A pesar de los muchos intentos de Fodor para explicar el significado de esta tesis, me confieso incapaz de captar lo que involucra. Mi respuesta no es distinta a la de las primeras reacciones a esta afirmación: parece que argumentar que todos (o casi) nuestros conceptos son innatos es una reducción al absurdo de la hipótesis del lenguaje del pensamiento. La tesis del innatismo de los conceptos va claramente más allá de la idea de que debemos tener un mecanismo gracias al cual los conceptos que *adquirimos* sean conceptos que *podemos* adquirir. Lo que de hecho mantiene es que, aparte de una minoría de conceptos composicionales, los conceptos que adquirimos son conceptos que ya tenemos. Y su defensa de un modelo atomista de la posesión de conceptos lo sitúa en la difícil situación de tener que responder algunas preguntas bastantes extrañas, tales como ¿cuántos conceptos tenemos?, ¿cómo se relacionan unos con otros?, ¿cómo se relacionan

---

<sup>53</sup> Fodor, “*The Present Status...*” cit.

<sup>54</sup> Fodor, “Fixation and Concept Acquisition” y “Reply to Putnam”, en M. Piattelli-Palmarini (ed.), *Language and Learning*, London, Routledge, 1980, p.151.

con los conceptos públicos?, o ¿cómo es posible que poseamos innatamente el concepto de “quark”?

Sólo la asunción de una teoría atomista de la adquisición de conceptos (*building-block theory*), de acuerdo con la cual primero adquirimos conceptos y sólo después aprendemos a hacer oraciones con ellos, involucra un compromiso tan implausible con el innatismo. La alternativa más obvia es una apelación más fuerte al lenguaje público para explicar la adquisición de conceptos y para evitar las dificultades de Fodor. Este tipo de teoría, más holista, puede todavía hacer uso de mecanismos discriminadores pre-lingüísticos (que tienen el papel de condiciones posibilitadoras de la adquisición de conceptos, pero no el papel constitutivo que Fodor quiere que jueguen), pero pone el énfasis en las conexiones inferenciales fundamentales que los conceptos deben tener con otros conceptos para merecer ser llamados tales. La perspectiva cambia si, en lugar de plantear la tarea de la psicología como algo posterior a una teoría del contenido, empezamos planteándonos qué son los conceptos (y, paralelamente, qué son las otras nociones semánticas tales como significado, referencia o verdad) dado el papel que juegan en nuestra comprensión de la mente. Para Fodor, los conceptos son entidades, cosas, que nuestra teoría descubre. Veremos que no es ésta la única concepción posible.

#### 4. *Externismo y contenido restringido*

##### 4.1. *Atomismo e internismo: un dilema*

Como hemos visto, Fodor mantiene una teoría representacional de la mente (4) y una teoría sintáctica o computacional de la mente (5). Es posible para él defender (4) y (5) simultáneamente porque:

- (a) las propiedades semánticas de un estado cognitivo sobrevienen en sus propiedades sintácticas, y
- (b) la interacción computacional entre características sintácticas sobreviene en procesos cerebrales.

Esta forma de sobreviniencia le permite mantener que las generalizaciones de sentido común hechas sobre la base de las actitudes proposicionales pueden ser analizadas de forma sistemática (tesis 6 y 7). La sobreviniencia de lo computacional en lo cerebral permite que alguien comprometido con el carácter nomológico de la causalidad (3b) y con el fisicismo (2) pueda también mantener que los estados mentales tienen poderes causales (3a) lo que implica realismo con respecto a lo intencional (1).

En esta sección me opondré a una consecuencia de (a) y (b):

(c) el contenido sobreviene en estados cerebrales (o, más generalmente, en estados internos del organismo).<sup>55</sup>

(c) es una variedad de lo que he llamado sobreviniencia estrecha al principio de este trabajo. Mi objeción a (c) hará uso de conocidos experimentos de pensamiento ofrecidos por Putnam y Burge. De acuerdo con Fodor, el contenido de los estados mentales es parte de sus condiciones de identidad, pero estos experimentos muestran que el contenido de los estados mentales no sobreviene (estrechamente) en estados del organismo al que son adscritos. Esta dificultad obliga a Fodor a elegir entre:

(d) proponer una noción de contenido tal que el contenido simultáneamente sirva para identificar estados cognoscitivos y sobrevenga en estados internos del organismo, o

(e) dar sentido a la idea de que el contenido sobreviene (ampliamente) en estados del organismo más el medio y la comunidad lingüística,

---

<sup>55</sup> La sobreviniencia es una relación transitiva. Si M sobreviene en C y C en F, entonces M sobreviene en F: si no puede haber cambios en M sin cambios en C y no puede haber cambios en C sin cambios en F, no puede haber cambios en M sin cambios en F.

donde ambas opciones tienen que respetar la que podemos llamar ‘condición atomista’, (f), i.e.:

- (f) no abandonar la posibilidad de proporcionar condiciones de identidad atomistas para los estados mentales.

Esta sección intenta mostrar que la opción (d) no es accesible (es decir, el contenido restringido no es contenido), que si lo fuera no podría respetar la condición atomista (f), y que la reacción atomista de Fodor (la ‘teoría de la dependencia asimétrica’) tampoco puede evitar el holismo. La sección 5 intenta mostrar que (e) también conduce al holismo, o bien al esencialismo y a la idea de armonía preestablecida entre mente y mundo.

#### *4.2. Externismo semántico y psicológico*

Los estados mentales, tal y como los individúa el sentido común, no sobrevienen en la constitución física de la persona a la que se adscriben porque esa constitución física es sólo un aspecto de su personalidad. Esto puede defenderse tanto como una posición de sentido común como filosóficamente (la filosofía de Strawson es un ejemplo excelente de lo segundo). En esta sección revisaré los argumentos de Putnam y Burge contra una concepción internista del contenido mental. La naturalización de la semántica se compromete, como mínimo, con la idea de que los estados mentales sobrevienen en estados físicos. Argumentaré que el naturalismo puede, a lo sumo, aspirar a una teoría del significado que tome como la base en la que sobreviene el significado al individuo *más* su medio (físico y social).<sup>56</sup>

La distinción misma entre internismo y externismo es fuente de confusión. El uso de “externismo” para designar el rechazo de la idea de que el contenido mental es una cuestión “interna” a la mente parece aceptar que la mente tiene un

<sup>56</sup> Cf. Heil, J., *The Nature of True Minds*, Cambridge, Cambridge University Press, 1992, para una clara enunciación de esta posición. Ver también para un comentario de la misma: Pinedo, M. de, “Review of John Heil’s *The Nature of True Minds*”, *International Journal of Philosophical Studies* 4, (1996).

dentro y un fuera. De hecho, a menudo esta aceptación es explícita: por ejemplo, Putnam y Burge, al menos en sus escritos de los setenta que comento en esta sección, argumentan contra la identificación de los estados mentales haciendo referencia exclusiva a lo que hay dentro del cuerpo, sin cuestionar la existencia de una “frontera” entre la mente y la realidad; Fodor en su desplazamiento en la dirección del externalismo en *The Elm and the Expert*<sup>57</sup> también retiene el compromiso con la concepción de la mente como algo con fronteras. Tomaré partido con Putnam y Burge en sus afirmaciones negativas acerca de las condiciones de identidad de los estados mentales, pero también argumentaré que no cabe encontrar sentido en una noción de contenido que separa un factor interno (‘restringido’) de otro externo (‘amplio’). La idea de que no cabe una separación de principio entre la contribución del sujeto y la del mundo a los estados mentales del primero tiene su origen en la filosofía kantiana y permite disolver algunas de las perplejidades que rodean a la naturalización de la semántica.

Según Putnam, las explicaciones tradicionales del significado descansan en dos asunciones no cuestionadas: “(I) Que conocer el significado de un término sólo consiste en estar en un cierto estado psicológico [...] [y] (II) que el significado de un término (en el sentido de ‘intensión’) determina su extensión (en el sentido en el que equivalencia en la intensión implica equivalencia en la extensión)”.<sup>58</sup> La idea de Putnam es, como veremos, que conocer el significado de un término no consiste *sólo* en estar en un cierto estado psicológico. Su argumento va contra una consecuencia de mantener (I) y (II) simultáneamente:

...dos hablantes no pueden estar en el mismo estado psicológico en todos los aspectos y comprender el término A de forma diferente; el estado psicológico del hablante determina la intensión [porque el estado psicológico estrecho “saber el significado de A” significa

---

<sup>57</sup> Cf. Fodor, *The Elm and the Expert*, Cambridge, Mass., The MIT Press, 1994.

<sup>58</sup> Putnam, “The Meaning of ‘Meaning’”, en *Mind, Language and...* cit., p. 219.

conocer la intensión de A, por la asunción (I)] (y por tanto, por la asunción (II), la extensión) de A.<sup>59</sup>

De las asunciones (I) y (II) se sigue: (III) los estados psicológicos determinan las extensiones. Sin embargo, podemos mostrar que (III) es falso. Putnam hace esto pidiéndonos que nos imaginemos la siguiente situación: hay un Tierra Gemela que es exactamente igual a la Tierra excepto que en lugar de agua los habitantes están rodeados de otro líquido, de idéntica apariencia y uso, la composición del cual no es H<sub>2</sub>O sino otra, que abreviamos como XYZ. Supongamos que no sé nada acerca de la química del agua (y, por supuesto, tampoco sabe nada mi gemelo acerca de la composición del agua gemela, dado que somos copias exactas); todo lo que sabemos es que es un líquido, que sale del grifo, que quita la sed, etc. Cuando pensamos “quisiera un vaso de agua” o cuando decimos “este agua está sucia”, de hecho nuestros pensamientos y preferencias son sobre sustancias distintas, los míos sobre *agua* (que es H<sub>2</sub>O) los de mi gemelo sobre *agua-gemela* (que es XYZ). Lo que sabemos sobre “agua”, nuestras intensiones, no determina las extensiones de nuestros conceptos. Si el significado se identifica con la intensión, entonces el significado no determina la extensión. Si el significado incluye a la extensión, entonces no está en la cabeza, sino que incluye las características pertinentes del medio. El ejemplo de Burge, como veremos, extiende esta inclusión desde el medio físico hasta la comunidad de hablantes.

Putnam explica el sentido en el que usa la noción de ‘estado psicológico’ introduciendo el concepto de ‘solipsismo metodológico’. Dice que el solipsismo metodológico asume que “ningún estado psicológico, propiamente hablando, presupone la existencia de individuo alguno que no sea el sujeto al que el estado se adscribe”.<sup>60</sup> Esta idea es explícita en la filosofía de Descartes e implícita en toda una tradición de la filosofía de la psicología (hemos visto como Fodor la acepta ex-

---

<sup>59</sup> *Ibid.*, p. 222.

<sup>60</sup> *Ibid.*, p. 220.

plícitamente). Putnam distingue entre ‘los estados psicológicos en sentido amplio’ (que se entienden apelando a las entidades acerca de las cuales versan) y ‘estados psicológicos en sentido restringido’ (que son permitidos por el solipsismo metodológico), y mantiene que los programas que abrazan el solipsismo metodológico han fracasado a lo largo de los tres últimos siglos. La tesis de Putnam es que la extensión *no* está determinada por los estados psicológicos y que, por tanto, tenemos que abandonar (I) o (II). Para él es (I) la problemática, ya que el significado puede definirse fácilmente de forma tal que incluya la extensión (la segunda parte de “The meaning of ‘meaning’”, del que estoy citando, tiene como objetivo desarrollar esta propuesta, pero no me detendré en ello).

(I) es la tesis central de lo que Burge ha llamado ‘individualismo’ y su rechazo es también la base de la premisa central de la crítica de Burge a la ciencia cognitiva: el significado es algo más que lo que sobreviene en los estados físicos internos de los organismos:

Las concepciones individualistas actuales de los estados y eventos mentales intencionales han tendido a adoptar dos formas. Una de ellas mantiene que el que un individuo esté en cualquier estado intencional [...] puede ser *explicado* haciendo referencia a estados y eventos de ese individuo que pueden ser especificables sin usar ningún vocabulario intencional y sin presuponer nada acerca del entorno social y físico del sujeto [...]. La otra forma de individualismo se sigue de la primera, pero es más débil. No intenta explicar nada. Sencillamente hace una afirmación de *sobrevenida*: los estados y eventos intencionales (tipos y ejemplares) no podrían ser distintos de lo que son, dadas las historias física, química, neural y funcional del individuo, donde estas historias se especifican no intencionalmente y de una forma independiente de las condiciones físicas y sociales exteriores al cuerpo del individuo.<sup>61</sup>

La tesis individualista entraña que la identidad de estados cerebrales es condición suficiente para la identidad de actitudes proposicionales, pero la identidad de estados cerebrales

---

<sup>61</sup> Burge, “Individualism and Psychology”, *The Philosophical Review* 95, (1986), p. 4.

no puede ser una condición suficiente para la identidad de actitudes proposicionales porque las actitudes proposicionales se individualizan con respecto a sus propiedades relacionales (es decir, con respecto a propiedades para las que el medio y la comunidad lingüística son pertinentes). Dos cerebros idénticos compartirían sus propiedades psicológicas sólo si estuvieran situados en el mismo medio y “perteneccieran” a la misma comunidad lingüística. El externismo es la alternativa de Burge al individualismo, y mantiene que el contenido de las actitudes proposicionales no sobreviene en ningún estado especificable no intencionalmente; las características sociales y ambientales son, en esta concepción, psicológicamente relevantes.

Para probar su tesis anti-individualista, Burge nos propone un experimento mental. El primer paso describe a una persona (Mary) que tiene muchas actitudes proposicionales cuya expresión incluye la palabra “artritis” y que afirma sufrir de artritis y que también piensa que tiene artritis en el muslo. El segundo paso consiste en una suposición contrafáctica: todo es igual que lo descrito en el primer paso, excepto algo que hace referencia al ambiente social de Mary-gemela. En el mundo real “artritis” sólo se aplica a dolencias en las articulaciones, pero en mundo gemelo los médicos, lexicógrafos y otras personas informadas utilizan “artritis” para referirse no sólo a la *artritis* sino también a otras dolencias reumatoideas. El veredicto de Burge es que en el caso contrafáctico, Mary-gemela carece de las actitudes que en el mundo de Mary serían atribuidas por medio del contenido de oraciones que contengan “artritis”. Mary y su gemela son físicamente idénticas, pero mientras el concepto de “artritis” de Mary se refiere a enfermedades de las articulaciones (es decir, a la *artritis*) y ella se equivoca al pensar que sufre esta enfermedad en su muslo, el concepto de “artritis” de Mary-gemela se refiere a otra cosa (a *artritis-gemela*) y acierta al pensar que la sufre en su muslo. La diferencia entre los contenidos de sus estados mentales es atribuible a diferencias en sus entornos sociales. En suma, las experiencias de las pacientes, sus esta-

dos y eventos fisiológicos, y cualquier secuencia de estados (no intencionalmente descritos) que medien entre su input y su output, permanecen constantes, mientras que el contenido de sus actitudes es distinto.

Burge extiende su experimento y sus conclusiones a muchos otros casos, algunos de los cuales dependen de una comprensión parcial de un concepto más que de un error acerca de él. Hace esto porque “[l]o que he llamado ‘comprensión parcial’ es común e incluso normal en el caso de un gran número de expresiones de nuestros vocabularios”.<sup>62</sup> El mismo tema reaparece en el artículo de Burge “Other bodies”,<sup>63</sup> donde se enfatiza que podemos sacar las mismas consecuencias anti-individualistas de cambios en el medio físico<sup>64</sup> que de cambios en el uso de las palabras en la comunidad lingüística.<sup>65</sup>

#### 4.3. *El contenido restringido y el solipsismo metodológico*

La ciencia cognitiva, como Fodor la entiende, tiene en las actitudes proposicionales su principal instrumento teórico. Por otro lado, la ciencia debe dar explicaciones causales de aquello que está causalmente relacionado en la naturaleza<sup>66</sup> o, lo que viene a ser lo mismo, debe ofrecer generalizaciones causales en virtud de las propiedades causales intrínsecas de las cosas. Sin embargo, el contenido de las actitudes proposicionales es, al menos, parte de sus condiciones de identidad. Y, finalmente, como Putnam y Burge intentan demostrar, es improbable que los contenidos sobrevengan en características físicas del cerebro. La fuerza de las consideraciones externalistas nos obliga a abandonar (1) [realismo intencional], (2) [físicismo], (3a) [causación mental] o (4) [carácter nomológico de la causalidad], o varios de ellos. O, alternativamente,

---

<sup>62</sup> *Ibid.*, p. 83 (He suprimido las cursivas de Burge).

<sup>63</sup> Cf. Burge, “Other Bodies”, en A. Woodfield (ed.) *Thought and Object*, Oxford, Clarendon Press, 1982, pp. 97–120.

<sup>64</sup> Cf. Putnam, “The Meaning of ‘Meaning’”, en *Mind, Language and...* cit., 1975, pp. 215–271.

<sup>65</sup> Cf. Burge, “Individualism and the Mental”, *Midwest Studies in Philosophy* 4, (1979), pp. 73–121.

<sup>66</sup> Fodor, “*Psychosemantics...*” cit., p. 34.

apelar a una concepción nueva, técnica, del contenido e intentar mantener todas estas tesis. O, finalmente, abrazar el externismo semántico y argumentar que no conduce al externismo psicológico.

Voy a revisar una variedad de respuestas dadas por Fodor para bloquear la necesidad de abandonar el fisicismo y el carácter nomológico de la causalidad. A diferencia del realismo intencional y de la afirmación de que existe causación mental, las tesis (2) y (3b) son conceptuales más que empíricas. Cabría argumentar que, frente a lo que algunos filósofos parecen asumir, dado que no tienen el estatuto de tesis científicas sino más bien de tesis a priori acerca de la naturaleza del conocimiento científico y de su papel en nuestra visión del mundo, son incompatibles con la idea de que tiene que existir un marco racional de justificación previo para que la ciencia pueda aspirar a presentar sus resultados como conocimiento correcto o verdadero. Sin embargo, en esta sección y en la siguiente de lo que me ocuparé será de los argumentos empíricos de Fodor contra las conclusiones sacadas por parte de Putnam y Burge de los casos de tierras gemelas y mostraré que la noción alternativa de contenido presentada por Fodor no puede hacer lo que él espera de ella.

Incluso concediéndole a Fodor que los estados intencionales juegan su papel no en virtud de su contenido sino en virtud de sus características formales, esta idea depende de que tales características “imitan” las características semánticas. Pero si las características semánticas dependen del medio físico y cultural, necesitamos una noción de contenido diferente que sea tan interna como la organización sintáctica de los estados mentales, esto es, que no dependa de nada externo a la piel del organismo. Mi respuesta rápida a la utilización del ‘contenido restringido’, el contenido de los estados mentales que es independiente del medio y sólo depende de la contribución del organismo (es decir, el contenido que podría sobrevenir en estados físicos internos), es que, o bien no es una noción de contenido, o bien una vez que se suprime la contribución externa a los estados mentales, sus condiciones

de identidad dependen de toda la red de actitudes proposicionales. El rechazo del contenido restringido y del internismo involucra abandonar (6) [sintacticidad de la mente] y (10) [solipsismo metodológico] que son tesis auxiliares que apoyan el paso de (4) [teoría representacional de la mente] y (5) [teoría computacional de la mente] a (8) [lenguaje del pensamiento].

Aunque inicialmente Fodor considera que el contenido restringido sobreviene en propiedades puramente formales de los estados mentales,<sup>67</sup> pronto, en “Cognitive science and the Twin–Earth problem”,<sup>68</sup> hace explícito que el contenido restringido no puede ser sólo sintaxis. En este artículo ofrece la propuesta, desarrollada plenamente en “Banish DisContent”<sup>69</sup> y *Psychosemantics*,<sup>70</sup> de que el contenido restringido es una función de contextos a condiciones de verdad. Con ello intenta mostrar que el argumento contra la idea de que variables psicológicas determinan el contenido de las representaciones mentales no tiene la fuerza que se le supone:

Necesitamos una noción de contenido de una representación mental para reconstruir la noción de contenido de una actitud proposicional *de dicto*, y necesitamos una noción de actitud proposicional *de dicto* para reconstruir la noción de la intencionalidad de la conducta; y necesitamos la noción de intencionalidad de la conducta para establecer una serie de generalizaciones psicológicas que parecen ser (más o menos) sostén de contrafácticos y verdaderas, y que subsumen a la conducta en virtud de su satisfacción de descripciones intencionales.<sup>71</sup>

Si la ciencia cognitiva debe usar propiedades intencionales de las representaciones mentales en la explicación de la conducta, entonces “no podemos salvar el programa de la

---

<sup>67</sup> Cf. Fodor, “Methodological Solipsism Considered as a Research Strategy in Cognitive Science”, en *Representations...* cit., pp. 225–253.

<sup>68</sup> Cf. Fodor, “Cognitive Science and the Twin Earth Problem”, *Notre Dame Journal of Formal Logic* 23, (1982), pp. 98–118.

<sup>69</sup> Cf. Fodor, “Banish DisContent”, en W. Lycan (ed.), *Mind and Cognition*, Oxford, Blackwell, 1990, (1986), pp.420–438.

<sup>70</sup> Fodor, “*Psychosemantics...*” cit.

<sup>71</sup> Fodor, “*Cognitive Science and...*” cit., p. 101.

ciencia cognitiva sintácticamente”.<sup>72</sup> El método de Fodor para salvar la ciencia cognitiva de las objeciones semánticas de Putnam y Burge es mantener que los significados son construcciones lógicas a partir de las actitudes proposicionales *de dicto* de los hablantes y oyentes. “[L]a gente normalmente conoce sus propias intenciones *de dicto*; las actitudes proposicionales *de dicto* están “en la cabeza” incluso si los significados no lo están”.<sup>73</sup> Aquí Fodor usa la distinción entre actitudes proposicionales especificadas *de dicto* y *de re* para reflejar la distinción entre estados psicológicos en sentido restringido y en sentido amplio. Con ello quiere mostrarse que incluso si personas molecularmente idénticas pueden hablar lenguajes distintos (es decir, sus palabras y conceptos pueden tener distintos significados dados distintos contextos), esto no significa que puedan diferir en sus actitudes proposicionales *de dicto*.<sup>74</sup> En resumen, Fodor argumenta que el contenido de una creencia no puede identificarse con sus condiciones de verdad y que para pasar del primero a las segundas necesitamos especificar un universo de discurso para las variables ligadas. Esta idea parte de la sugerencia de que “las intenciones son funciones de *contextos* y pensamientos a condiciones de verdad”.<sup>75</sup> Mis contenidos son idénticos a los de mi gemelo aunque sus condiciones de verdad no lo sean, dado que vivimos en contextos diferentes. Analizando el contenido de esta forma tenemos un criterio extensional para el contenido restringido.

Una solución más elaborada puede encontrarse en el capítulo 2 de *Psychosemantics*. Fodor acepta que las historias de Putnam y Burge muestran, como mínimo, que no podemos

---

<sup>72</sup> *Ibid.*, p. 102.

<sup>73</sup> *Ibid.*, p. 103.

<sup>74</sup> Sin embargo, este paso podría también cuestionarse. Existen argumentos poderosos a favor de la idea de que las actitudes *de re* son más fundamentales que, y condición de posibilidad de, las actitudes *de dicto*, y que por ello estas últimas están infectadas de la dependencia contextual de las primeras. Ver Burge, “Belief *De Re*”, *Journal of Philosophy* 74, (1977), pp. 338–362 y también Evans, G., *The Varieties of Reference*, Oxford, Clarendon Press, 1982; edited by John McDowell.

<sup>75</sup> Fodor, “*Banish DisContent...*” cit., p. 424.

esperar que una ciencia respetable se apoye en actitudes individuadas como lo hace el sentido común.<sup>76</sup> Necesitamos otras condiciones de identidad para los estados mentales distintas a las del sentido común. Sin embargo, mantiene que esto no pone el peligro el uso explicativo de las actitudes porque, una vez que se cambian los límites, sus propiedades semánticas y causales todavía serán adscribibles y los principios de individuación fáciles de reformular.

La conexión entre identidad extensional e identidad de contenido no se rompe (como Putnam y Burge mantienen haber mostrado), sólo se relativiza al contexto. Antes de “The meaning of ‘meaning’” los contenidos eran:

...algo como funciones de pensamientos a condiciones de verdad, [ahora] [...] los contenidos son funciones de contextos y pensamientos a condiciones de verdad, [es decir] [...] los contenidos todavía determinan las extensiones relativamente al contexto. [Por tanto] [...] los contenidos de dos pensamientos son idénticos si aplican la misma relación desde pensamientos y contextos a condiciones de verdad.<sup>77</sup>

Lo que Fodor propone es un criterio extensional para contenidos restringidos: mis pensamientos y los de mi gemelo tienen el mismo contenido pero distintas condiciones de verdad porque ocurren en distintos contextos. El contenido amplio de un pensamiento se obtiene cuando se especifica el restringido y se fija el contexto.

#### 4.4. ¿Es el contenido restringido autónomo?

No. El contenido restringido puede ser un instrumento útil para una teoría *computacional* de la mente, pero no tiene utilidad para explicar el carácter *representacional* de la mente. En esta subsección llamo la atención sobre dos aspectos problemáticos del proyecto fodoriano de aplicar la noción de contenido restringido para explicar la intencionalidad de la mente. Primero argumentaré que es dudoso que sea conteni-

---

<sup>76</sup> Fodor, “*Psychosemantics...*” cit., p. 30.

<sup>77</sup> *Ibid.*, pp. 47-48.

do, para a continuación poner en cuestión que las condiciones de identidad del contenido restringido sean atomistas. En la siguiente sección discutiré la teoría de la dependencia asimétrica, la propuesta de Fodor para complementar al contenido restringido en la explicación de la intencionalidad.

*¿Por qué el contenido restringido no es contenido?*

No es difícil responder a esta pregunta. Si una teoría del contenido es una teoría del significado y los significados son evaluables semánticamente, entonces no puede obtenerse una teoría completa del significado porque los contenidos restringidos, por definición, sólo son evaluables semánticamente si un contexto se ha fijado. Y para fijar un contexto necesitamos un mundo. Pero dado que el contenido restringido es una respuesta a la supuesta necesidad de adoptar el solipsismo metodológico, cuando hablamos de contenido restringido no hablamos del mundo. Los contenidos restringidos son sólo contenidos potenciales. Para convertirlos en contenido genuino algo más es necesario.

*¿Son las condiciones de identidad de los contenidos restringidos atomistas?*

A no ser que el contenido restringido esté anclado de alguna forma al mundo, sus condiciones de identidad sólo pueden establecerse en términos del papel que juega en la red funcional/computacional en la que opera. Y, a no ser que pueda distinguirse entre aquellas relaciones que sean relevantes para su identidad y aquellas que no, parece que todas las conexiones son necesarias para identificar un contenido restringido particular. Pero esto sería circular: necesitamos que el contenido restringido pueda identificarse atomísticamente si queremos que la intencionalidad sea capturada por redes nomológicas que expliquen la causación de la conducta. No podemos apelar a las redes funcionales de representaciones mentales para individuar el contenido restringido porque primero necesitamos alguna forma de individuación para poder hablar de una red funcional.

Esta dificultad ha sido señalada por varios filósofos. Stalnaker, por ejemplo, afirma que dado que:

...el medio externo no puede ser pertinente [y] las diferencias *físicas* internas que no se reflejen en diferencias en la organización funcional tampoco pueden ser pertinentes [...] las condiciones de identidad de los contenidos restringidos deben estar *totalmente* determinadas de una forma u otra por las relaciones funcionales entre los ejemplares que poseen el contenido y otros eventos, estados y objetos mentales y, posiblemente, otros ejemplares.<sup>78</sup>

Block<sup>79</sup> también lanza esta crítica haciendo uso de un ejemplo del artículo de Putnam “Computational psychology and interpretation theory”. Dos niños viven en distintas partes de un país y la única diferencia lingüística importante entre las comunidades donde viven es que la palabra “grug” se refiere a la plata en una y al aluminio en otra. Ambos niños saben algunas cosas sobre el grug pero no suficiente para fijar la referencia (supongamos que ambos saben las mismas cosas: “las cacerolas están hechas de grug”, etc.). Si nos limitamos a la noción restringida de contenido, los conceptos de grug de ambos niños tienen el mismo contenido. Los niños nunca llegan a conocerse ni a descubrir la diferencia lingüística entre ambas comunidades. Cuando crecen aprenden español y a traducir la palabra respectivamente por “plata” y “aluminio”. El contenido restringido de sus conceptos es ahora diferente como consecuencia de un pequeño cambio en la información colateral, concretamente haber aprendido cómo traducir la palabra a una lengua extranjera.<sup>80</sup> Cualquier intento de separar las diferencias objetivas en la extensión de la noción de contenido restringido de una representación mental (es decir, de separar la extensión del contenido restringi-

---

<sup>78</sup> Stalnaker, R., “How to do Semantics for the Language of Thought”, en B. Loewer y G. Rey (eds.) *Meaning in Mind (Fodor and His Critics)*, Oxford, Blackwell, 1991, p. 232.

<sup>79</sup> Block, N., “What Narrow Content Is Not”, en B. Loewer y G. Rey (eds.) *Meaning in Mind (Fodor and His Critics)*, Oxford, Blackwell, 1991, pp. 33–64.

<sup>80</sup> Putnam, “Computational Psychology and Interpretation Theory”, *Realism and Reason. Philosophical Papers, vol.3*, Cambridge, Cambridge University Press, 1983, pp.144–147.

do) tendrá que enfrentarse al hecho de que una palabra o un concepto se ven continuamente modificados por la adquisición de información colateral. El contenido restringido, por sí mismo, es holista.

La situación en la que nos encontramos es una en la que el contenido restringido necesita algo más aparte del lugar que ocupa dentro de la red funcional para que sea contenido y para tener condiciones de identidad atomistas. La siguiente sección repasará dos propuestas de Fodor a este respecto. La primera intenta conectar contenidos restringidos a propiedades mundanas atomistas, y parece fracasar. La segunda abandona completamente la noción de contenido restringido pero a cambio abraza una variedad de la armonía preestablecida difícil de digerir.

#### 4.5. Una formulación inicial del holismo del significado

Putnam<sup>81</sup> atribuye a Dummett<sup>82</sup> la acuñación del término ‘holismo del significado’ y ofrece una definición muy clara de esta teoría:

Tal y como ha sido desarrollada por Davidson (a partir de doctrinas de Quine), la doctrina de Holismo del Significado mantiene que, aún siendo cierto que un hablante debe haber aprendido una gran cantidad de oraciones y de procedimientos de fijación de creencias conectadas con esas oraciones antes de que quepa hablar de ‘significados’ en relación a sus preferencias, no existe un conjunto individual definido de tales oraciones que uno *deba* ser capaz de usar para poseer un ‘concepto’ particular. Más bien es una cuestión de interpretación. El intérprete implícitamente construye una ‘teoría del significado’ que asigna una interpretación particular a cada palabra, por ejemplo, ‘oro’ o ‘silla’, parcialmente sobre la base del uso del hablante de oraciones que contienen esa palabra, pero también sobre la base de los significados asignados a las *otras palabras de esas oraciones por esa misma teoría del significado*, y sobre la base de las situaciones en la que esas *otras* palabras son usadas. La interpretación, según el holista del significado, es una cuestión de nego-

<sup>81</sup> Putnam, “Models and Modules”, *Cognition* 17, 1984, pp. 253-264.

<sup>82</sup> Dummett, M., “What is a Theory of Meaning”, en *The Seas of Language*, Oxford, Clarendon Press, 1993, (1975), pp. 1-33; y ver también: Dummett, “What is a Theory of Meaning (II)”, en *The Seas of Language*, Oxford, Clarendon Press, 1993, (1976), pp. 34-93.

ciaciones (*trade off*), y esas negociaciones “cuando son racionales son pragmáticas”, en frase de Quine. No existen reglas precisas de interpretación.<sup>83</sup>

La conclusión que Putnam saca es que las ciencias cognitivas no pueden hacer nada con una teoría del significado porque: “la teoría de la interpretación y la ciencia cognitiva se ocupan de proyectos muy diferentes [...]”.<sup>84</sup> “La explicación en términos de creencias y deseos pertenece al nivel que he estado llamando *teoría de la interpretación*”.<sup>85</sup>

El argumento de Quine a favor del holismo en “Dos dogmas”<sup>86</sup> va dirigido contra la teoría del significado positivista: los significados de oraciones individuales no son entidades mentales o platónicas. Debemos reconocer que lo que se enfrenta al tribunal de la experiencia es el cuerpo de oraciones (o el sistema completo de doctrina en desarrollo). La propuesta de Quine es que las oraciones son significativas sólo en tanto que contribuyen al funcionamiento de todo el lenguaje (esta propuesta sigue la idea fregeana de que las palabras sólo tienen significado cuando contribuyen sistemáticamente a las condiciones de verdad de oraciones completas).

Putnam distingue entre el holismo con respecto al significado y el holismo con respecto a la fijación de creencias (esta distinción no es explícita en Quine). Hay dos formas distintas de negar el holismo semántico: (i) negar el holismo en general (incluso con respecto a la fijación de creencias); y (ii) aceptar la explicación holista de Quine de la fijación de creencias pero mantener que podemos conseguir una ontología científica o filosófica de significados.

La primera línea de oposición al holismo (tomada, por ejemplo, por Dummett)<sup>87</sup> podría aceptar que las consideraciones de Quine son correctas en la ciencia teórica. Los significados de términos teóricos dependerían de la teoría donde

---

<sup>83</sup> Putnam, “*Models and Modules...*”, cit., p. 257.

<sup>84</sup> Putnam, “*Computational Psychology and...*” cit., p.150.

<sup>85</sup> *Ibid.*, p. 154.

<sup>86</sup> Cf. Quine, W., “Two Dogmas of Empiricism”, en *From a Logical Point of View*, Cambridge, Mass., Harvard University Press, 1953, pp. 20–46.

<sup>87</sup> Cf. Dummett, “*What is a...*” cit.

aparecen. Esto lleva a una compartimentación del lenguaje. Sin embargo, distinguir entre el uso ordinario y el científico de un término parece una política errónea y poco económica. No hay razones para negar que nuestro lenguaje es una aventura cooperativa y que el lenguaje científico y el cotidiano son diferentes pero interdependientes.<sup>88</sup>

La alternativa a esta línea de argumentación contra el holismo semántico [(ii)] es poner en duda el paso desde el holismo con respecto a la fijación de creencias hasta el holismo del significado. Esto se logra defendiendo que los significados son entidades internas:

Estas entidades estarían, de alguna forma, *asociadas con palabras, morfemas y oraciones concretos* (de manera que los significados serían *no-holistas*). Por otro lado, el carácter holista de la *fijación de la creencia* se explica *postulando que el paso del significado de las oraciones a sus condiciones de afirmabilidad involucra un "procesamiento de arriba a abajo"*; esto es, el uso general de la inteligencia y de la información disponible.<sup>89</sup>

¿Qué importancia puede tener el lenguaje del pensamiento para la noción de significado? Según Putnam, si existe, su carácter debe ser como el de cualquier otro lenguaje. El carácter holista de la fijación de creencias es un requisito de la presunción de racionalidad necesaria para comprender lo mental. ¿Qué razón existe para pensar que existe una noción precisa, independiente del contexto y más que meramente pragmática, de igualdad de significado para el lenguaje del pensamiento si no podemos encontrarla o construirla para otros lenguajes?

Las representaciones mentales cambian en función del cambio en los estereotipos que el sujeto asocia a ellas y, por tanto, "la igualdad o diferencia de representación mental sería un criterio neurofisiológico útil para la igualdad o diferencia de estereotipo pero [...] los estereotipos no son *signifi-*

<sup>88</sup> Putnam, "Meaning Holism", en L.E. Hahn y P. Schlipp (eds.), *The Philosophy of W. V. Quine*, La Salle, Open Court, 1986, pp. 408-409.

<sup>89</sup> *Ibid.*, p. 409.

*cados*”.<sup>90</sup> Esto es así porque, incluso si hay representaciones mentales con contenido, los contenidos no son internos. Ningún tipo de entidad interna puede corresponderse con las relaciones semánticas, dado que las primeras son extremadamente variables bajo los procesos de fijación de creencias mientras que las segundas no lo son.

*5. De la dependencia asimétrica al contenido amplio: el problema del holismo semántico*

Ya tenemos todas las piezas (menos una) para terminar el puzzle de Fodor: el pensamiento es manipulación computacional de representaciones compuestas de ‘conceptos restringidos innatos’. Las representaciones tratan del mundo porque los contenidos restringidos se ven disparados por propiedades de las cosas que hay en el mundo (estos procesos de disparo son la pieza que falta; hablaré de ellos en esta sección). La manipulación computacional es posible porque las propiedades sintácticas de las representaciones reflejan sus propiedades semánticas; y los “conceptos restringidos” explican la intencionalidad porque están, *de hecho*, conectados a estados del mundo y, por tanto, tienen contenido. Pero, retomando nuestras preguntas de la sección anterior: ¿es el contenido restringido contenido? Y, ¿pueden darse condiciones de identidad atomistas una vez que el medio (físico y social) se excluye?

Hemos visto que la respuesta intuitiva a la primera pregunta es un claro “no”. Incluso si el contenido restringido se ancla al mundo por medio de algún mecanismo causal, no es nada claro que los contenidos restringidos por sí mismos sean suficientes para que los procesos computacionales instanciados en el cerebro sean procesos ejecutados sobre estados *intencionales*. Por otra parte, si no podemos apelar al mundo para identificar los contenidos de las representaciones mentales, parece que sus condiciones de identidad dependerán sólo de su lugar dentro de la red funcional de estados mentales, lo que es incompatible con el atomismo. Es el momento para

---

<sup>90</sup> *Ibid.*, p. 421.

escuchar lo que Fodor tiene para responder a estas preocupaciones.

### *5.1. Anclando el contenido restringido al mundo: la dependencia asimétrica*

Los estados mentales con contenido tienen propiedades semánticas. Ahora bien, ¿en virtud de qué conexiones con el mundo se obtienen estas propiedades? Fodor comienza considerando la que llama ‘teoría causal tosca’, de acuerdo con la cual “un símbolo expresa una propiedad si es nomológicamente necesario que *todas* las instancias de la propiedad y *sólo* ellas causan ejemplificaciones (*tokenings*) del símbolo”,<sup>91</sup> donde la causación ha de entenderse contrafácticamente. Sin embargo existen problemas obvios tanto con la condición “todo” como con la condición “sólo”. Para empezar, ¿cómo podemos explicar la representación errónea (*misrepresentation*), es decir, los fallos en la cláusula “sólo”?

De acuerdo con Fodor, la demolición por parte de Chomsky de la teoría de la adquisición del lenguaje de Skinner deja intacta la propuesta semántica de este último: que una palabra “w” exprese una propiedad, *ser un w*, se reduce a ciertas disposiciones conductuales bajo ciertos estímulos. Esta propuesta es naturalista ya que no hace uso en su explicación del significado de nociones semánticas o intencionales, y es atomista.

Se puede aceptar una semántica skinneriana para el pensamiento (el cual, según Fodor y a diferencia de la conducta lingüística, no es un tipo de comportamiento voluntario) aún rechazando su teoría del lenguaje.

El hecho de que este estado exprese esta propiedad se reduce al hecho de que ejemplificaciones del estado sean, en el sentido pertinente, respuestas discriminadoras a instancias de la propiedad; es decir, las instanciaciones del estado covarian con (están ‘bajo el control de’) instanciaciones de la propiedad, y esta covariación es legaliforme y, por tanto, respetuosa con los contrafácticos.<sup>92</sup>

<sup>91</sup> Fodor, “*Psychosemantics...*” cit., p. 100.

<sup>92</sup> Fodor, “A Theory of Content, I: The Problem”, en *A Theory of Content and Other Essays*, Cambridge, Mass., The MIT Press, 1990, pp. 55–56.

Skinner sin su conductismo y sin su teoría del aprendizaje nos deja con una teoría como la ‘semántica informacional’ de Dretske en su *Knowledge and the Flow of Information*.

La base de la teoría de Dretske es que “[e]ventos-S (por ejemplo, ejemplificaciones de símbolos) expresan [llevan información, diría Dretske] la propiedad P si la generalización ‘Ps causan Ss’ es contrafáctica (*counterfactual supporting*)”.<sup>93</sup> Dretske exige que ‘Ps y sólo Ps causan Ss’ sea una ley. El contenido de un símbolo, en esta teoría, está determinado plenamente por relaciones nómicas (esto es, por las *disposiciones* a ejemplificarlos) y es, por tanto, independiente de la historia de sus ejemplificaciones *de hecho*.

De acuerdo con Dretske, dos eventos están en una relación semántica si uno “transmite información acerca del otro” y la relación depende de la correlación entre ellos en el proceso de aprendizaje. La representación errónea puede explicarse señalando que las ejemplificaciones de los símbolos están causadas por el evento equivocado (siendo los eventos correctos los que están sistemáticamente relacionados con el símbolo en el proceso de su aprendizaje).<sup>94</sup> Un problema obvio con esta propuesta es que la distinción entre el periodo de aprendizaje y lo que ocurre después no es estricta y además, señala Fodor, esta solución sólo se aplicaría a símbolos aprendidos, no a la información innata. Sin embargo, la principal dificultad la plantean los contrafácticos: si se da una correlación errónea durante el periodo de aprendizaje y el evento equivocado dispara un ejemplar del símbolo, el evento sería parte de la información (el significado) transmitido por el símbolo y no podríamos hablar de representación errónea. Apelar al refuerzo no puede ayudar, como mostró Chomsky: “lo que se refuerza es una cosa, lo que se aprende es frecuentemente otra”.<sup>95</sup>

---

<sup>93</sup> *Ibid.*, p. 57 (He suprimido el énfasis).

<sup>94</sup> Cf. Dretske, *Knowledge and the...* cit.

<sup>95</sup> Fodor, “*A Theory of...*” cit., p. 63.

Una propuesta distinta, también descartada por Fodor, es la de Stampe.<sup>96</sup> Stampe subraya las diferencias contrafácticas entre representaciones correctas y “desbocadas” (*wild*): el camino causal correcto es el único que produce ejemplares correctos en ‘circunstancias óptimas’. Estas circunstancias deben ser definidas en términos no semánticos y no intencionales para evitar la circularidad. Stampe hace esto apelando a la teleología, es decir, al hecho de que los mecanismos de fijación de creencias funcionan como se supone que tienen que hacerlo.

Estos intentos de resolver el problema de la representación errónea se basan en la estrategia general de distinguir entre dos tipos de situaciones de tal forma que la covariación nómica determina el significado en una pero no en la otra. Al primer tipo de situación se la considera *Normal*. Los casos de error son casos en los que algo ha fallado. Lo que necesitamos es una explicación naturalista de la Normalidad. Una propuesta influyente se expresa en términos de “teleología (natural), es decir una noción más o menos darwinista/histórica de mecanismos biológicos *haciendo aquello por lo que han sido seleccionados*”.<sup>97</sup> Esto no significa, argumenta Fodor, que pueda haber una historia teleológica que contar acerca de la creencia, sino que hay una para “los mecanismos de fijación de las creencias”, esto es, no necesitamos hablar de la función de los estados cognitivos mismos. Por ello, el proyecto no implica que las creencias y los otros estados intencionales puedan individuarse en términos de sus funciones. Millikan, la más influyente defensora de este enfoque, parece pensar de otra forma; de acuerdo con ella los estados cognoscitivos se distinguen por sus funciones y es la función de un estado cognoscitivo la que determina su objeto intencional.<sup>98</sup> Fodor encuentra serias dificultades en definir cuál es la función propia de, digamos, un deseo. El contenido de un deseo no puede identificarse con aque-

---

<sup>96</sup> Cf. Stampe, “*Towards a Causal...*”cit.

<sup>97</sup> Fodor, “*A Theory of...*”cit., p. 64.

<sup>98</sup> Cf. Millikan, R., *Language, Thought and Other Biological Categories*, Cambridge, Mass., The MIT Press, 1984, pp. 55-64; citado en Fodor, “*A Theory of...*”cit., pp. 66–67.

llo que el deseo se supone que ayudará a causar, ni siquiera probabilísticamente, porque los deseos pueden fácilmente ayudar a causar algo diferente de lo que intuitivamente pensamos que es su contenido (por ejemplo, mi deseo de beber agua puede causar que beba ginebra en ciertas, poco deseables, circunstancias). Por otra parte, no hay garantía de que los mecanismos estén de hecho “diseñados” para conducir siempre a la verdad.<sup>99</sup> Y, en cualquier caso, no está claro cómo evitar el peligro de circularidad. No puede sencillamente asumirse que todos los mecanismos de fijación de creencias funcionen de la misma forma: discriminamos entre distintos tipos de mecanismos de fijación de creencias en términos de su contenido.

Las explicaciones teleológicas, según Fodor, no solucionan el ‘problema de la disyunción’ sino que meramente lo obvian. Este problema puede enunciarse como sigue: si *cada* ejemplar de un símbolo está causado por algo con una condición suficiente para causarlo, y cada propiedad cuya instanciación es suficiente para causar ejemplificaciones de un símbolo es expresada por ese símbolo (es su significado), entonces *cada* ejemplar del símbolo es causado por algo que pertenece a su extensión; pero entonces, la extensión de un símbolo incluye cualquier cosa que de hecho dispare ejemplares del mismo sin dejar espacio para la posibilidad del error. Los gatos serían parte de la extensión de “lince”, por ejemplo. Como contemos la historia teleológica acerca del contenido de un estado mental dependerá de qué descripción intencional del contenido nos parezca correcta porque podemos contar diferentes, e igualmente plausibles, historias naturalistas sobre el mismo símbolo. Por ejemplo, podemos justificar tanto que las ranas “se representan” moscas como que se representan puntos negros en movimiento. La razón para esto es que de hecho la mayoría de los puntos negros en movimiento que la rana se encuentra (y sus ancestros encontraron) eran moscas. “[A] Darwin no le importa cómo describas los objetos intenciona-

---

<sup>99</sup> Para este tipo de crítica a la semántica teleológica, ver por ejemplo, Stich, S., *From Folk Psychology to Cognitive Science: The Case Against Belief*, Cambridge, The MIT Press, 1983.

les a los que las ranas lanzan la lengua. [...] *A Darwin le importa cuántas moscas te comes, no bajo que descripción lo haces*".<sup>100</sup> "En la práctica, las explicaciones darwinistas tratan a las representaciones fiablemente coextensas como sinónimas: mientras que, por supuesto, la explicación psicológica no lo hace".<sup>101</sup> Seguimos con el mismo problema que empezamos, "la indeterminación intencional masiva".<sup>102</sup> El diagnóstico de Fodor se expresa, una vez más, en términos de contrafácticos. Un intento darwinista de resolver el problema de la disyunción sólo tiene en cuenta las historias de selección *de hecho*, mientras que para la adscripción intencional debemos tener en cuenta situaciones contrafácticas.

La solución ofrecida por Fodor en *Psychosemantics* comparte con la semántica teleológica el uso de propiedades contrafácticas, aunque apela a propiedades diferentes. La idea que sirve de base para la propuesta es que las falsedades dependen ontológicamente de las verdades (sólo podemos tener creencias falsas sobre cosas acerca de las cuales tenemos creencias verdaderas): "En particular, identificar erróneamente una vaca como un caballo no me llevaría a decir 'caballo' *excepto cuando existiera independientemente una relación semántica entre ejemplificaciones de 'caballo' y caballos*".<sup>103</sup> La relación causal entre vacas y ejemplificaciones de "caballo" 'depende asimétricamente' de relaciones causales entre caballos y ejemplificaciones de "caballo".

Los problemas con la condición "todos" de la teoría causal (necesitamos que todas las instancias de la propiedad causal en ejemplificaciones del símbolo) son más profundos que los

<sup>100</sup> Fodor, "A Theory of..." cit., p. 72-73.

<sup>101</sup> *Ibid.*, p. 78.

<sup>102</sup> *Ibid.*, p. 75. Otra dificultad con las opciones teleológicas digna de mencionar es que, incluso en circunstancias normales, una ejemplificación de un símbolo puede ser producida por algo que no es lo que representa sin que sea el caso que ha habido algún error. Mis pensamientos sobre ranas pueden disparar pensamientos sobre moscas sin que haya ninguna a mi alrededor, pero el contenido de "mosca" no es pensamientos-raniles. No hay nada anormal acerca de estos procesos, pero la teoría teleológica no puede separar *moscas de pensamientos-raniles* como contenidos de "mosca".

<sup>103</sup> Fodor, "Psychosemantics..." cit., p. 107.

problemas con la cláusula “sólo” (necesitamos que sólo las instancias de la propiedad...), como el de la representación falsa que hemos discutido. ¿Qué clase de contrafáctico podría explicar la relación causal sin usar vocabulario semántico o intencional? La respuesta de Fodor depende de la existencia bruta de mecanismos fiables de correlación entre mente y mundo, mecanismos que no es necesario poder especificar. “‘Caballo’ significa *caballo* si las ejemplificaciones de ‘caballo’ están fiablemente causadas por propiedades psicofísicas de las cuales las instanciaciones de *caballo* son de hecho causalmente responsables”.<sup>104</sup> Lo que hace que “caballo” signifique *caballo* es que ambos están conectados, no cómo estén conectados.

Existe un problema fundamental con el tipo de semántica denotacional que Fodor se ve obligado a desarrollar para dar fortaleza a su hipótesis del lenguaje del pensamiento. Al reducir el significado a denotación o referencia y al mantener que un concepto significa la propiedad que normalmente lo dispara, uno necesita explicar cómo hay que entender esas propiedades. El no hacerlo convierte en imposible saber cómo han de comprenderse los significados de los conceptos correlacionados con las propiedades. Una respuesta que ha ganado mucho apoyo recientemente es pensar acerca de las propiedades esencialmente, siguiendo a Kripke.<sup>105</sup> Para Fodor, el significado de “agua” es *agua*, y la propiedad *ser agua*, en la lectura esencialista, no es otra que *tener la composición H<sub>2</sub>O*. Aunque para Fodor *agua* y *H<sub>2</sub>O* sean propiedades diferentes que proporcionan el significado a diferentes conceptos, todas las cosas con la primera propiedad deben tener la segunda. Alguien puede poseer el concepto “agua” sin saber que *agua* es *H<sub>2</sub>O*. Sin embargo, si entendemos el significado de manera esencialista, necesitamos apelar a *H<sub>2</sub>O* para individuar la propiedad *agua* que es el significado de “agua”.

---

<sup>104</sup> *Ibid.*, p. 122.

<sup>105</sup> Cf. Kripke, S., “Naming and Necessity”, en D. Davidson y G. Harman (eds.) *Semantics of Natural Language*, Dordrecht, D.Reidel, 1972, pp. 253–355.

El esencialismo de Fodor es aún más notable en su formulación de la relación de dependencia asimétrica en términos de mundos posibles. “1. Aes causan ‘A’es; 2. Ejemplares de ‘A’ *no* están causados por Bes en mundos posibles cercanos en los que Aes *no* causan ‘A’es; 3. Aes causan ‘A’es en mundos posibles cercanos en los que Bes no causan ‘A’es”.<sup>106</sup> La dificultad es doble. Por un lado, se nos deja en la oscuridad acerca de cómo determinar qué mundos posibles son cercanos al nuestro y cuáles no. Cualquier apelación al contexto para determinar esto corre el peligro de ser circular. Si nuestra clasificación de los mundos posibles es sensible al contexto o al interés violaríamos el principio naturalista que exige dar una explicación del significado en términos no intencionales. Establecer qué mundos posibles son más cercanos no parece factible sin tener en cuenta qué es pertinente y qué no lo es en los contrafácticos utilizados para relacionar causalmente propiedades y representaciones mentales. Adaptando un ejemplo de Lewis<sup>107</sup> que llama la atención sobre la sensibilidad con respecto al contexto de las dependencias contrafácticas, necesitaríamos tener en cuenta cuáles son nuestros propósitos explicativos antes de decidir acerca de la verdad de estos dos enunciados: (I) Si Maradona hubiera sido inglés, habría sido jugador de cricket; (II) Si Maradona hubiera sido inglés, habría sido futbolista. Hay contextos en los que (I) puede usarse para expresar algo verdadero y contextos en los que puede usarse (II). No está claro que pueda evitarse mencionar el contexto de discurso para decidir qué relación de similitud es pertinente para la verdad del condicional (en nuestro ejemplo, si la relación entre Maradona y ser futbolista, o entre ser inglés y jugar al cricket): parece que necesitaríamos preguntar a la persona que enuncia los condicionales si está pensando en un Maradona anglicado o no. Lewis, por ejemplo, argumenta coherente a favor del realismo con respecto a los mundos posibles<sup>108</sup> reconociendo que la similitud

---

<sup>106</sup> Fodor, “*Psychosemantics...*” cit., p. 109.

<sup>107</sup> Lewis, D., *Counterfactuals*, Oxford, Blackwell, 1973, pp. 66-67.

<sup>108</sup> *Ibid.*, p. 84-91.

o cercanía entre estos mundos y el mundo real es relativa al interés.<sup>109</sup> Sin embargo, Fodor no puede aceptar el método de Lewis para su individuación<sup>110</sup> a no ser que pueda eliminar la dependencia del contexto y del interés de este método. Estas consideraciones llevan a adoptar, con Davidson, la idea de que los referentes son abstracciones de una totalidad, el mundo, al que nuestro lenguaje, como un todo, refiere. Para determinar qué entidades consideramos que existen en el mundo necesitamos una estructura compleja, una teoría. Y para entender cómo esta teoría determina el significado de sus términos tenemos que tener en cuenta todos los significados y sus relaciones estructurales. Por ello no cabe, a pesar de Fodor, dar una descripción de la relación entre símbolos y entidades en términos no semánticos y no intencionales.<sup>111</sup>

Es precisamente este tipo de alternativa holista acerca del significado la que Fodor intenta evitar con su propuesta causal. Sin embargo, no hace ningún intento de argumentar contra el carácter intencional (dada la dependencia del interés) de la determinación del valor de verdad de los contrafácticos. Fodor, y aquí es donde comienza la segunda dificultad, permanece igualmente silencioso con respecto a qué sostiene la presunción de que Bes no causan 'A's en mundos cercanos donde A's no causan 'A's. Según su análisis del significado, "vaca" significa vacas en vez de vacas o caballos porque (1) las vacas causan "vaca", (2) si las vacas no causaran "vaca" nada lo haría, y (3) incluso si los caballos no causaran "vaca", las vacas todavía lo causarían. De acuerdo con (2), nin-

---

<sup>109</sup> *Ibid.*, p. 91-95.

<sup>110</sup> "Si  $p$  fuera el caso, entonces  $q$  sería el caso" es verdadero en el mundo real si, y sólo si, hay un mundo- $p$  donde  $q$  es el caso más cercano al mundo real que cualquier mundo- $p$  donde  $q$  no es el caso". Véase *Ibid.*, p. 49.

<sup>111</sup> Debo esta crítica a discusiones con Argiri Arfani y Murali Ramachandran. Robert Brandom ofrece un argumento similar contra la tentación de usar el fiabilismo (*reliabilism*) como excusa epistemológica para el naturalismo: "los hechos con respecto a la fiabilidad de un mecanismo pueden ser objetivos [...] sólo [...] relativamente a una clase de referencia. [...] Pero la elección de la clase de referencia adecuada no está ella misma objetivamente determinada por hechos especificables con vocabulario naturalista" En Brandom, R., "Insights and Blindspots of Reliabilism", *The Monist* (81), 1998, pp. 383-384.

guna cosa, ni siquiera los caballos (que se asume que causan “vaca” en el mundo real), podría causar “vaca” si las vacas no lo hicieran. ¿Por qué es esto así? La respuesta, que Fodor no ofrece, tendría que ser naturalista, es decir, (2) sólo podría apoyarse en relaciones causales entre vaca y “vaca”. Nos quedamos con una forma de esencialismo sin sostén argumental que considera que representaciones y propiedades están causalmente conectadas. Y parece que tal esencialismo y tal análisis causal del significado depende de la idea de que las representaciones mentales y las características del mundo están relacionadas independientemente de cualquier cosa sabida o percibida por el portador de la representación. Criticaré esta idea en la próxima sección.

### *5.2. El contenido amplio y la armonía preestablecida*

Hasta aquí he discutido la defensa de Fodor de una noción de contenido restringido para retener tanto el carácter computacional como el intencional de las generalizaciones psicológicas. No obstante, Fodor ha saltado recientemente del contenido restringido al amplio, es decir, de una semántica internista a otra externista. Este salto se acepta explícitamente en *The Elm and the Expert*:

Solía pensar que el problema de construir [...] teorías que conectarán leyes intencionales con sus implementaciones computacionales simplemente *exigía* que las leyes intencionales se construyeran restringidamente. Pero ahora me inclino a pensar que no es así y que una semántica puramente informacional puede después de todo reconciliarse con una teoría de los procesos mentales puramente computacional.<sup>112</sup>

Este paso en la dirección del contenido amplio responde a una de las dificultades a las que el contenido restringido no podía enfrentarse, el carácter semántico del contenido. Pero Fodor todavía tiene que explicar cómo los contenidos amplios pueden ser utilizables en los procesos computacionales implementados en el cerebro. Recordemos, Fodor intenta hacer

---

<sup>112</sup> Fodor, *The Elm and...* cit., p. 17.

tres tesis, con las cuales ya estamos familiarizados, compatibles: (1) Necesitamos una teoría del contenido porque las generalizaciones de la psicología intencional seleccionan a las creencias y deseos que subsumen haciendo referencia a sus contenidos; (2) El contenido de un pensamiento depende de la forma en que se relaciona con el mundo y no de otros pensamientos y sus contenidos; y (3) Las leyes psicológicas están implementadas en procesos computacionales (sintácticos).

El contenido restringido daba una buena respuesta a (3) pero fracasaba con respecto a (2). El contenido amplio encaja con (2), pero ¿puede hacerse compatible con (3)? Lo que Fodor intenta mostrar es que las características importantes del contenido restringido pueden obtenerse también por medio de una individuación amplia del contenido de las actitudes proposicionales. Esta conexión entre contenido amplio y estrecho tiene que responder a dos tipos de críticas<sup>113</sup> que Fodor llama:

(i) ‘Casos tipo Putnam’: el mismo contenido restringido se corresponde con dos, o más, contenidos amplios distintos. Estos casos pueden dividirse en dos tipos:

(i.a) ‘Casos de gemelos’: ver los ejemplos de “agua” y “artritis” estudiados en la sección 4.2. Mi pensamiento “el agua es húmeda” y el de mi gemelo comparten características computacionales; sin embargo, su individuación amplia difiere;

(i.b) ‘Casos de conocimiento parcial’: el clásico ejemplo de Putnam es alguien que tiene tanto el concepto de “olmo” como el de “haya” sin ser capaz de distinguir olmos de hayas. Para una persona sin esa capacidad, y en lo que se refiere a sus mecanismos internos, ambos conceptos tienen las mismas

---

<sup>113</sup> Fodor considera tres, siendo la tercera el problema de la inescrutabilidad de la referencia. No obstante, una corta discusión de los dos primeros será suficiente para mis propósitos.

características; no obstante tienen distinto contenido amplio;<sup>114</sup>

(ii) ‘Casos tipo Frege’: dos, o más, contenidos restringidos se corresponden con el mismo contenido amplio. Por ejemplo, yo creo que “el lucero matutino es Venus” pero no que “el lucero vespertino es Venus”; es difícil explicar esto si lo único que tenemos para individuar el contenido de estos pensamientos es el referente de los conceptos que lo componen.

Veamos cómo responde Fodor a estos casos. Los casos de gemelos (i.a) se rechazan como sigue: las generalizaciones que capturan lo que hay en común entre mi gemelo y yo son generalizaciones accidentales; el hecho de que se escapen a la individuación amplia del contenido no pone en peligro el estatuto científico de la ciencia porque la ciencia no tiene que preocuparse de generalizaciones accidentales. No argumentaré contra esta respuesta.

Los casos de conocimiento parcial se responden señalando que la semántica, en la concepción denotacional de Fodor, se ocupa fundamentalmente de contrafácticos. Por ejemplo, en el caso ‘olmo–haya’ mencionado lo que es importante con respecto a la individuación del contenido amplio no es si yo de hecho puedo o no distinguir entre las condiciones de aplicación de los conceptos “olmo” y “haya”, sino “si *podría* distinguirlos si me importara”<sup>115</sup> consultando a un experto. Cómo evitar que esta apelación a un experto reintroduzca el holismo (¿cómo, por ejemplo, se determina quién es y quién no un experto sin recurrir a creencias que conducen a otras creencias y así sucesivamente?) no es algo de lo que Fodor se ocupe. Pero la mayor consecuencia de separar lo que es accesible al sujeto de lo que podría ser accesible es que, para dar sentido al externismo, debe mantenerse que “el corazón del

<sup>114</sup> Sin embargo, y a la luz de la distinción entre externismo semántico y externismo psicológico que he defendido en la sección 4.2, cabría decir que alguien sin la capacidad de distinguir entre olmos y hayas no posee los conceptos (aunque puede usar las palabras confiando en que los expertos de la comunidad los doten de contenido).

<sup>115</sup> Fodor, *The Elm and...* cit., p. 37.

externismo [es] que *la semántica no es parte de la psicología*".<sup>116</sup> Esto es, los estados intencionales son amplios, pero ¡los procesos psicológicos son estrechos!

Con respecto a los casos fregeanos la idea es culpar al carácter *ceteris paribus* de las adscripción psicológica de estados por los casos aberrantes donde el contenido restringido y el amplio no coinciden. Según Fodor, las leyes físicas son las únicas sin excepciones, es decir, sólo la física está causalmente cerrada: no puede haber una fuerza causal que interfiera en una cadena de causación física y que no sea capturable por leyes físicas. Esto no es el caso en las otras ciencias. Por ejemplo, imaginemos que existe una ley biológica que dice que las proteínas y los hidratos de carbono producen grasas saturadas al combinarse. No se considerará un contraejemplo a esta ley si ocho minutos antes de que me coma un bocadillo de jamón el sol se convierte en supernova aunque la cadena causal (biológica) se rompe y no hay leyes biológicas que expliquen esto.<sup>117</sup> La física es la reina y las ciencias especiales son sus súbditas. Los procesos psicológicos instancian leyes psicológicas, a no ser que hechos físicos (distintos de los que realizan los procesos psicológicos, se supone) se metan en medio. Hay una dificultad muy seria con esta estrategia: para que funcione ha de ser el caso que los casos anormales no proliferen, de manera que puedan ser tratados como excepciones y no como contraejemplos. Pero, para bloquear esta proliferación, necesitamos apelar a una tesis metafísica muy fuerte, una especie de armonía preestablecida entre las características sintácticas de los contenidos de nuestros estados mentales y la estructura real del mundo: "es *plausible* [...] que el mundo esté organizado de tal forma que prohíba la proliferación de casos de gemelos y de casos tipo Frege". "La armonía así establecida entre el contenido amplio y el papel computacional, aunque sea contingente, es suficientemente

---

<sup>116</sup> *Ibid.*, p. 38.

<sup>117</sup> Cf. Fodor, "Special Sciences", en *Representations*, Brighton, Harvester Press, 1981, (1974), pp.127–145. También ver, Fodor, "Propositional Attitudes..." cit. pp.177–203.

fiable".<sup>118</sup> No parece posible establecer cuántos casos de gemelos hay y, en cualquier caso, lo importante no es su número sino su mera posibilidad. Sin embargo, los casos de Frege *si* que proliferan y es patente que lo hacen de forma ilimitada: no hay, por ejemplo, dos o veinte formas de pensar sobre Venus, como muestra el número potencialmente infinito de descripciones imaginables para referir a ese planeta. El mundo es tan rico y complejo que siempre está abierta la posibilidad para que nuevos conceptos surjan para entenderlo. Sin embargo, para Fodor, los conceptos son innatos y cortan al mundo por sus articulaciones: la estructura del mundo y la estructura de nuestro bagaje conceptual coinciden y es por ello por lo que partes del mundo tienen éxito en activar partes de nuestras mentes. Cuando no lo tienen es o accidental o irrelevante para la explicación de la conducta. El único proceso de adquisición de conceptos aceptado por Fodor es la activación de los conceptos que el sujeto ya posee innatamente. Y, para evitar que tal repertorio conceptual esté desconectado de cómo es el mundo, hay que *asumir* que las características del mundo que activan nuestros conceptos innatos se corresponden a contenidos estrechos de manera *anterior* e *independiente* de tales activaciones. Parece claro que tal coincidencia sería demasiado extraña para ser correcta.

Claro está, la relación entre pensamientos y estados de cosas no es mera coincidencia, pero parece haber una forma más económica de explicarlo. Nuestra caracterización del mundo y nuestra caracterización de la mente no se hacen usando vocabularios distintos, sino que son ambas resultado de nuestros esfuerzos por dar sentido a la interacción racional de las criaturas con su medio. Fodor, por el contrario, está comprometido con un realismo metafísico fuerte que considera al mundo como algo fijado de una vez por todas (es decir, las verdades sobre el mundo ya están todas en su sitio antes de que nuestra investigación empiece, siendo nuestra investigación un mero intento de descubrir tantas de estas verdades como sea posible) y confía en que nuestro aparato

---

<sup>118</sup> Fodor, *The Elm and...* cit., pp. 28 y 54.

cognoscitivo filtro, de alguna forma, al mundo de manera fiel  
a la estructura preorganizada de éste.

Departamento de Filosofía, Universidad de Granada (España)  
Programa “Ramón y Cajal”, Ministerio de Ciencia y Tecnología  
e-mail: [manueldpg@hotmail.com](mailto:manueldpg@hotmail.com)