

MODULARIDAD Y TEORÍA COMPUTACIONAL DE LA MENTE EN LA OBRA DE JERRY FODOR.

Nota crítica en torno a *The Mind Doesn't Work that Way*¹

PABLO RYCHTER

Universidad de Buenos Aires y Universitat de Barcelona

Desde finales de la década de 1960, Jerry Fodor ha hecho importantes contribuciones a ámbitos disciplinares diversos: filosofía, lingüística y psicología cognitiva. Si bien el espíritu naturalista con el que Fodor se mueve permite y justifica los saltos de un dominio a otro, parece inevitable que su influencia en cada uno de ellos haya sido asimétrica. En el ámbito filosófico, Fodor es conocido como un defensor de la teoría representacional de la mente (Fodor (1975) y (1987)) que en los últimos años ha decidido librar una larga batalla en favor del atomismo semántico (Fodor (1987), (1990), (1998)). Por otro lado, en el ámbito de las ciencias cognitivas propiamente dichas, Fodor es más famoso por sus tesis acerca de la arquitectura mental, más específicamente por su elucidación de la noción de modularidad (Fodor (1983)). A lo largo de los años, el mismo Fodor ha mantenido esta bifurcación temática y son excepcionales las ocasiones en las que sus trabajos más “filosóficos” hacen mención a la tesis de la modularidad (y a la inversa, en sus trabajos acerca de la modularidad, rara vez se ocupa de problemas semánticos o filosóficos). La publicación de su más reciente libro² es una buena oportunidad para preguntarnos por la relación entre las distintas facetas de Fodor, en particular por la relación y coherencia entre sus puntos de vista acerca de la modularidad y su defensa de la teoría representacional de la mente. Intentaré mostrar que entre estas dos líneas de investigación hay una tensión de difícil resolución. Veremos que en su último libro Fodor retoma una línea de pensamiento

¹ Una versión preliminar de este trabajo fue presentada en el seminario interno del proyecto UBACYT F003 “Categorías, conceptos y significado”. Agradezco a los miembros de dicho proyecto por los comentarios recibidos.

² Fodor (2000), *The Mind Doesn't Work that Way: The Scope and Limits of Computational Psychology*, Cambridge Mass, MIT Press. Salvo indicación en sentido contrario, las citas y números de página corresponden a esta obra.

pesimista (presentada por primera vez en el apéndice de *El lenguaje del pensamiento*³ y formulada en sus términos actuales en *La modularidad de la mente* bajo el título de “primera ley de Fodor sobre la inexistencia de la ciencia cognitiva”) difícilmente conciliable con la manera en que él mismo presenta y defiende la teoría representacional de la mente y la hipótesis del lenguaje del pensamiento. Me ocuparé de este problema en la segunda sección del presente trabajo. Previamente, la sección 1 estará dedicada a presentar los temas que constituyen el núcleo del nuevo libro de Fodor, para lo cual será necesario remontarnos a obras anteriores, especialmente a Fodor (1983).

1. Cómo no funciona la mente

Pese a sus austeras dimensiones (sólo 126 páginas), el libro de Fodor es notoriamente denso y cubre una amplia variedad de temas. Como su título lo indica, constituye en gran parte una respuesta a los puntos de vista defendidos por Pinker (1997) y Plotkin (1997), quienes junto con L. Cosmides y J. Tooby han gestado lo que Fodor llama “nueva síntesis”, una integración de la teoría computacional de la mente (TCM) con temas propios de la psicología evolucionista. Pero, por otro lado, el lector encontrará también desarrollos independientes (acerca de la modularidad, el nativismo, la psicología “racionalista”, etc.) que explicitan o completan ideas previamente presentadas por Fodor de manera más programática y menos acabada.

Si bien son varios los excesos que Fodor reprocha a la nueva síntesis, la discusión se centra en tres temas centrales: el alcance desmedido que se le concede a la TCM, la injustificada adopción de la tesis de la modularidad masiva y el coqueteo con la psicología evolucionista. Comenzaremos nuestro comentario por el primer punto, que es para Fodor el objetivo central del libro: “echar luz sobre la cuestión de qué es lo que está bien y qué es lo que está mal en la idea de que la mente es una computadora” (p. 3).

A) Los límites de la TCM

La TCM es una de las hipótesis básicas y fundacionales de la ciencia cognitiva clásica. Su núcleo es la idea según la cual hay un

³ Agradezco al evaluador anónimo de *Análisis Filosófico* el señalamiento de este punto histórico.

nivel de descripción en el que los procesos mentales pueden ser vistos como *computacionales*. La noción de computación supuesta por la TCM en sus presentaciones estándar⁴ es la noción clásica, según la cual computar es efectuar transformaciones de símbolos sólo en virtud de las propiedades sintácticas de sus constituyentes. Estas operaciones sintácticas están implementadas o “cableadas” (*hard-wired*) en el cerebro, de la misma manera que las operaciones más básicas de un ordenador están implementadas en su *hardware*. Vale la pena recordar que Fodor ha sido siempre uno de los mayores publicistas de esta idea, a la que frecuentemente se refiere como “la enseñanza de Turing”⁵. Sin embargo, frente a los excesos entusiastas de Pinker y Plotkin, Fodor prefiere en este libro adoptar un tono de cautela y enfatizar los problemas con los que la TCM se encuentra a la hora de modelar los aspectos centrales de la cognición. La conclusión a la que arribará es que, por buena que sea la TCM, “no puede ser más que un fragmento de una psicología cognitiva completa y satisfactoria; y que los problemas más interesantes acerca del pensamiento probablemente no serán iluminados por ningún tipo de teoría computacional que hoy día podamos imaginar” (p. 1). Esta línea de pensamiento pesimista ya había sido presentada por Fodor en las últimas páginas de *La modularidad de la mente*, como epílogo a una imagen de lo mental cuyos lineamientos generales conviene recordar, ya que son el trasfondo supuesto por la discusión presente.

Según la imagen ofrecida por Fodor (1983) y (2000), la mente es un conjunto de subsistemas, cada uno de los cuales procesa información de manera relativamente autónoma. Estos subsistemas son clasificados en dos grupos: sistemas modulares y sistemas no modulares, según presenten o no una serie de propiedades características que Fodor enumera y especifica. Dentro de las propiedades que caracterizan a los sistemas modulares, hay dos que resultan definitivas o esenciales para la modularidad: el encapsulamiento informacional (cada módulo opera de manera autónoma y no tiene acceso a la información procesada por otros sistemas) y la especificidad de dominio (cada módulo está diseñado para dar respuesta a un problema específico y su capacidad de procesamiento no se extiende a

⁴ Cf. p. ej. Fodor (2000), pp. 3-4; (1998), p. 28; (1987), pp. 40-42; Pinker (1997), pp. 66-69.

⁵ Cf. p. ej. Fodor (1998), p. 28; (1994), pp. 9 y 14.

otros dominios)⁶. Una vez caracterizada de esta manera la noción de “sistema modular”, es una cuestión empírica la de establecer cuáles son los sistemas que *de hecho* la satisfacen. Hay dos grupos de sistemas que según Fodor resultan claramente modulares: los “sistemas de entrada” (cuya función es extraer información del mundo y codificarla simbólicamente de manera que pueda ser utilizada como *input* de un proceso computacional) y los “sistemas motores” (responsables de comportamientos como el habla y la locomoción). Entre los sistemas modulares mejor estudiados se encuentran los encargados de la percepción y análisis de estímulos lingüísticos, la percepción visual de objetos medianos, el reconocimiento de rostros y voces, etc. Por otro lado, hay sistemas que –también como una cuestión de hecho establecida empíricamente– *no* son modulares. Se trata de los llamados “sistemas centrales”, cuya función es recibir y elaborar la información proporcionada por los distintos sistemas de entrada, junto con la que está almacenada en la memoria. Más específicamente, sobre la base de dicha información los sistemas centrales se ocupan de la “fijación de creencias” mediante inferencias no-demostrativas. En términos de Fodor, los sistemas centrales son aquellos “a los que hacemos referencia cuando hablamos, en sentido preteórico, de procesos mentales como el pensamiento y la solución de problemas”⁷. Aunque primarios para nuestros intereses, son precisamente éstos los fenómenos que según el último Fodor quedan al margen de lo que puede ser explicado por “cualquier tipo de teoría computacional que hoy día podamos imaginar” (p. 1).

Las razones de este pesimismo también habían sido adelantadas en *La modularidad de la mente*. Allí Fodor señala que las inferencias propias de los sistemas centrales presentan dos características distintivas a las que llama “*isotropía*” y “*quineanismo*”. Por un lado, una inferencia es *isotrópica* cuando la información relevante para la conclusión puede provenir de cualquier ámbito de conocimiento previamente establecido. Fodor considera que un caso patente de isotropía se da en las inferencias propias de la práctica científica, tanto en el caso de la confirmación como en el del descubrimiento: “todo lo que el científico conozca puede servir para deter-

⁶ Véase la tercera parte de Fodor (1983) para un listado completo de las propiedades características de la modularidad. Por otro lado, el capítulo 4 de Fodor (2000) comienza con una interesante serie de precisiones acerca de la noción de módulo, incluyendo las diferencias con el uso del término que hace Chomsky.

⁷ Fodor (1983) p. 146.

minar, en principio, lo que ha de creer en lo sucesivo. Así, la botánica puede influir sobre la astronomía, siempre que pueda concebirse alguna conexión entre ambas⁸. Otro caso claro de isotropía se da en el razonamiento analógico corriente, “proceso que depende de la transmisión de información entre dominios cognitivos que en principio se suponían irrelevantes el uno para el otro”⁹. Por otro lado, una inferencia es *quineana* cuando está determinada por propiedades *globales u holísticas* del sistema de creencias. La explicitación detallada de esta noción de “*determinación global*” es sin duda uno de los principales aportes del nuevo libro de Fodor. En el capítulo 2 se trata esta cuestión y se la ejemplifica con la propiedad de la simplicidad: suele admitirse que entre dos creencias rivales (o entre dos cursos de acción rivales) es razonable optar por la más simple, *ceteris paribus*. Pero ocurre que la mayor o menor simplicidad de una creencia, argumenta Fodor, no depende únicamente de las propiedades intrínsecas de la creencia en cuestión (y en particular, no depende de la *sintaxis* de las representaciones mentales constitutivas de la creencia). Por el contrario, la simplicidad de una creencia depende del contexto en el que ésta se enmarque: la misma creencia puede contribuir de distinta manera a la simplicidad-complejidad de distintos corpus de creencias.

Ahora bien, la determinación global de los procesos centrales representa un problema para la TCM, problema que Fodor (1983) resumía en su “primera ley de Fodor sobre la inexistencia de la ciencia cognitiva”. Según esta ley, “cuanto más global (cuanto más isotrópico) es un proceso cognitivo, tanto menos se lo comprende. Los procesos *muy* globales, como el razonamiento analógico, no se comprenden en absoluto”¹⁰. Según Fodor, la ciencia cognitiva clásica (i. e., la que tiene la TCM como núcleo fundamental) no puede lidiar con propiedades globales como el quineanismo y la isotropía: los intentos de modelar computacionalmente los procesos que exhiben estas propiedades han conducido invariablemente al fracaso. La noción clásica de computación (i. e., la noción de transformación de símbolos en virtud de sus propiedades sintácticas) es un instrumento efectivo para reconstruir aquellos procesos que están determinados sólo de manera local, pero resulta insuficiente para modelar procesos que –como la inferencia abductiva de sentido común– pue-

⁸ Fodor (1983) p. 148.

⁹ Fodor (1983) p. 150.

¹⁰ Fodor (1985) p. 151.

den ser afectados por información proveniente de cualquier parte del corpus de creencias. Un ejemplo de esta dificultad es proporcionado por lo que en el ámbito de la inteligencia artificial se conoce con el nombre de "*frame problem*".

Sin embargo, continua el razonamiento pesimista de Fodor, no tenemos en la actualidad ninguna idea plausible acerca de cómo funciona la mente, que no sea la provista por la ciencia cognitiva clásica. Por tanto, aquellos procesos y capacidades mentales que resulten altamente isotrópicos y quineanos no son comprendidos ni lo serán en tanto no surja una alternativa plausible a la ciencia cognitiva basada en la TCM. Vale la pena recalcar que este pesimismo alcanza a aquellos procesos mentales que comúnmente y de manera paradigmática llamamos "pensamiento", dado que son estos procesos los que alcanzan niveles máximos de determinación global: "cuando más nos acercamos a lo que desde una perspectiva preteórica consideramos como los ejercicios 'superiores', 'más inteligentes', menos reflejos y menos rutinarios de las capacidades cognitivas, tanto más tienden a manifestarse en ellos propiedades globales como la isotropía"¹¹. Los procesos modulares, por el contrario, están exentos de dichas propiedades globales y por tanto sí pueden ser comprendidos por la ciencia cognitiva clásica. De hecho, lo que a Fodor le interesa enfatizar en su último libro es que los procesos modulares constituyen el dominio al cual se limita nuestra posibilidad de comprender científicamente a la mente: "la teoría computacional es probablemente verdadera a lo sumo únicamente de las partes modulares de la mente", mientras que el funcionamiento de las partes no modulares "es, en términos de Chomsky, un misterio más que un problema" (p. 99). En estas condiciones, la propuesta metodológica de Fodor es la siguiente: la ciencia cognitiva debe concentrarse en el estudio de los procesos modulares, hasta tanto no surjan las ideas clave (probablemente ajenas al marco provisto por la TCM) que permitan el abordaje de los procesos no modulares: "en la actualidad, nos faltan algunas ideas fundamentales acerca de la cognición (...). No hay por qué lamentarse de esta situación. Sin duda alguien tendrá esas ideas tarde o temprano, y habrá entonces algún progreso. Hasta entonces, creo que seguimos un buen consejo si nos concentramos en aquellos problemas acerca de la mente sobre los que sabemos cómo pensar. Felizmente, parece que hay partes de la mente interesantes (aunque periféricas) que son modulares, aun

¹¹ Fodor (1983) p. 151.

cuando hay otras partes, *más interesantes* y menos periféricas, que no lo son. Y también parece que Turing estaba en lo cierto acerca de cómo pensar acerca de cómo funcionan las partes modulares. De manera que tenemos muchas cosas acerca de las cuales sabemos más o menos qué hacer” (pp. 99-100, énfasis agregado). Esta propuesta metodológica resulta paradójica y algo decepcionante: según ella, la ciencia cognitiva clásica (la única, por cierto, a la cual se le atribuye algún tipo de plausibilidad) debe *renunciar* a comprender los fenómenos centrales y más interesantes acerca de la mente (la cognición propiamente dicha, según algunos) y *limitarse* al estudio de aspectos relativamente periféricos (aunque no por ello carentes de interés, valga la aclaración). Como veremos en la sección 2, el carácter paradójico de la propuesta en cuestión se acrecienta cuando se la considera en el contexto de los argumentos con los que el mismo Fodor presenta y defiende la teoría representacional de la mente y la hipótesis del lenguaje del pensamiento.

B) La tesis de la modularidad masiva y la psicología evolucionista

Según hemos visto, aunque la TCM no pueda explicar el funcionamiento de los sistemas no modulares, sí constituye una teoría satisfactoria acerca de las partes modulares de la mente: la cognición modular es el tipo de proceso para el cual funciona bien el modelo clásico de computación. Por tanto, los problemas de la TCM se disolverían si fuera cierto que todas o la mayor parte de los procesos cognitivos son modulares. Esto es precisamente lo que afirma la tesis de la modularidad masiva (TMM) y lo que constituye el segundo eje de la discusión que Fodor entabla con la nueva síntesis.

Conviene tener en cuenta que, *prima facie*, la verdad de la TMM es una posibilidad coherente que podría establecerse empíricamente: recordemos que era una cuestión *fáctica* la de establecer cuáles sistemas son modulares y cuáles no. De hecho, buena parte de la presunta evidencia en favor de la TMM está constituida por argumentos empíricos de carácter adaptacionista. La contraevidencia presentada por Fodor, en cambio, es de carácter mixto: si bien se detiene a cuestionar la validez de diversos argumentos adaptacionistas (pp. 64-71), el núcleo de su ataque a la TMM está dado por un argumento de carácter conceptual que explota lo que él llama “*input problem*” (pp. 71-78). Según este argumento, *no es posible* que la mente conste sólo de sistemas modulares, dado que el funcionamiento de dichos sistemas supone que, como paso previo, un siste-

ma de dominio general haya clasificado a los diversos *inputs* como pertenecientes a distintos dominios y, por tanto, como correspondientes a distintos módulos. Si bien se trata de un argumento conceptual *a priori*, Fodor se preocupa por señalar que el *input problem* no es una mera ocurrencia filosófica, sino que tiene ejemplificaciones concretas en ciencia cognitiva: ni siquiera en el muy estudiado caso de la percepción del lenguaje está claro –para Fodor– cuál es el mecanismo por el cual el módulo correspondiente decide que un determinado *input* forma parte de su dominio. Por otro lado, la estrategia de Fodor para cuestionar los argumentos adaptacionistas a favor de la TMM consiste en mostrar que dichos argumentos, en caso de ser válidos, no muestran el carácter modular de un determinado mecanismo, sino sólo su carácter innato.

Pero además de rechazar los argumentos adaptacionistas específicamente presentados a favor de la TMM, Fodor completa su crítica a la nueva síntesis con un último capítulo dedicado a una revisión general de la psicología evolucionista. La tesis central de la psicología evolucionista afirma que “la arquitectura cognitiva es una adaptación evolutiva” y que, para comprender por qué la mente funciona como lo hace, debemos embarcarnos en una investigación histórica acerca de sus orígenes evolutivos. La evaluación que Fodor hace de los argumentos en apoyo de esta tesis es completamente negativa: no hay razones de fuerza por las que un innatista deba comprometerse con la psicología evolucionista. Más interesante aún, no hay razones de fuerza por las que un *naturalista* acerca de lo mental deba hacerlo. Rechazar la tesis central de la psicología evolucionista no nos obliga a enfrentarnos con “la imagen científica del mundo”, ni siquiera al “estudio biológico de la mente” (p. 90). En relación a este punto, Fodor señala que la politización de los temas relacionados con la evolución humana ha contribuido a la sobreestimación de los argumentos y a la aceptación apresurada de las tesis evolucionistas. Hay razones independientes para dudar de la psicología evolucionista, razones perfectamente respetables desde un punto de vista naturalista y que nada tienen que ver con el deseo de salvaguardar el “lugar del hombre en la creación”.

Por último, es interesante enfatizar la manera armónica en que se relacionan los distintos componentes de la nueva síntesis: TCM extendida a todos los sistemas de la mente, modularidad masiva y psicología evolucionista. Por un lado, la TMM hace posible la aplicación irrestricta de la TCM. Si, como afirma la TMM, todos los sistemas de la mente son modulares, entonces no hay procesos en

los que se manifieste de manera crítica la determinación global: en la medida en que un sistema está informacionalmente encapsulado (como, por definición, lo están los módulos), su funcionamiento está completamente determinado por la información disponible localmente y puede, por tanto, ser modelado sin mayores problemas mediante la noción clásica de computación. Si se adopta la TMM, entonces la TCM es una teoría acerca de los procesos mentales en general, y no sólo acerca de un subconjunto acotado de los mismos. Por otro lado, hemos dicho ya que la TMM está apoyada fundamentalmente por tesis provenientes de la psicología evolucionista: buena parte de las razones que se presentan a favor de la existencia de sistemas especializados para dominios específicos se basan en señalar la ventaja adaptativa que dichos sistemas representan para el organismo. De manera que al menos la incoherencia no es algo que se le pueda reprochar a la nueva síntesis.

2. Representación, computación y lenguaje del pensamiento

Volvamos ahora a nuestro problema inicial: cómo se concilian los puntos de vista de Fodor acerca de la modularidad (y en especial, su pesimismo acerca del alcance de la TCM) con los argumentos mediante los que defiende la teoría representacional de la mente (TRM) y la hipótesis del lenguaje del pensamiento (HLDP). Como he adelantado en la introducción del presente trabajo, mi impresión es que dicha conciliación no resulta sencilla. Expondré a continuación las razones de esta sospecha.

Veamos primero qué ocurre con la TRM. En *Conceptos*, Fodor comienza su presentación de la TRM reconociendo que dicha teoría es más bien “una vaga confederación de tesis a la que le falta una formulación canónica”¹². Sin pretensiones de subsanar esta falta, Fodor caracteriza a la TRM como la conjunción de las siguientes te-

¹² Fodor (1998), p. 24. La falta de una presentación canónica de la TRM y de sus relaciones con la TCM y la HLDP es una de las principales dificultades con las que nos encontramos aquí. Hay, por ejemplo, una notable diferencia en este respecto entre las presentaciones de Fodor (1998) y (1987), y a la vez entre éstas y exposiciones como la de Loewer y Rey (1991). A diferencia de estos autores, en este trabajo he optado por tratar a la TCM y a la TRM como tesis bien diferenciadas, ateniéndome a las definiciones para cada una de ellas que Fodor presenta en los textos citados en cada caso.

sis: (1) la explicación psicológica es típicamente nómica y completamente intencional, (2) las representaciones mentales son las portadoras primitivas de contenido intencional, (3) el pensamiento es computación, (4) el significado es información (más o menos), (5) lo que sea que distinga a conceptos coextensivos debe estar en la cabeza. Dado que para nuestros fines interesa la manera en que Fodor defiende la TRM más que la TRM en sí misma, las dos últimas tesis no son relevantes y nos concentraremos en la tercera: "el pensamiento es computación".

Ahora bien, cuando la TRM afirma que el pensamiento es computación, ¿a qué se refiere con "pensamiento"? (en otros términos, ¿qué partes de la mente deben ser tratadas como computacionales?). Una primera posibilidad es que por "pensamiento" se entienda sólo lo que ordinariamente llamamos de esa manera, i. e. la capacidad para resolver problemas complejos, la fijación de creencias, el razonamiento analógico, etc. Si este fuera el caso, la TRM estaría afirmando que los procesos mentales *centrales* son computacionales (lo cual, podrá advertirse ya, no es del todo coherente con el pesimismo de Fodor acerca de los alcances de la TCM). Una segunda posibilidad es que cuando se dice que "el pensamiento es computación", por "pensamiento" se entienda la vida mental en general, incluyendo no sólo los procesos que hemos llamado "centrales" sino también los periféricos: la percepción y los sistemas motores. Finalmente, una tercera posibilidad es que "pensamiento" se esté usando, de manera poco corriente, para hacer referencia sólo a los procesos periféricos. Si bien esta tercera opción es, según creo, la más consistente con los puntos de vista reconstruidos en la sección 1, claramente es inadecuada como una interpretación de lo que Fodor está afirmando aquí. Sólo las primeras dos opciones tienen alguna plausibilidad en este sentido, lo cual resulta evidente cuando se consideran las tesis (1) y (2) de la TRM. La manera en que Fodor elucida estas dos tesis nos permite ver con claridad cuáles son los estados y procesos acerca de los cuales versa la TRM. La tesis (1) implica que "los ejemplos paradigmáticos de leyes [psicológicas] son leyes acerca de las relaciones causales entre las creencias, deseos y acciones". Es decir, estados mentales propios de los sistemas centrales. Además, en la presentación de la tesis (2) Fodor introduce lo que ha sido siempre el núcleo de la TRM¹³: la idea según

¹³ Cf. p. ej. Fodor (1987), p. 43.

los que se manifieste de manera crítica la determinación global: en la medida en que un sistema está informacionalmente encapsulado (como, por definición, lo están los módulos), su funcionamiento está completamente determinado por la información disponible localmente y puede, por tanto, ser modelado sin mayores problemas mediante la noción clásica de computación. Si se adopta la TMM, entonces la TCM es una teoría acerca de los procesos mentales en general, y no sólo acerca de un subconjunto acotado de los mismos. Por otro lado, hemos dicho ya que la TMM está apoyada fundamentalmente por tesis provenientes de la psicología evolucionista: buena parte de las razones que se presentan a favor de la existencia de sistemas especializados para dominios específicos se basan en señalar la ventaja adaptativa que dichos sistemas representan para el organismo. De manera que al menos la incoherencia no es algo que se le pueda reprochar a la nueva síntesis.

2. Representación, computación y lenguaje del pensamiento

Volvamos ahora a nuestro problema inicial: cómo se concilian los puntos de vista de Fodor acerca de la modularidad (y en especial, su pesimismo acerca del alcance de la TCM) con los argumentos mediante los que defiende la teoría representacional de la mente (TRM) y la hipótesis del lenguaje del pensamiento (HLDP). Como he adelantado en la introducción del presente trabajo, mi impresión es que dicha conciliación no resulta sencilla. Expondré a continuación las razones de esta sospecha.

Veamos primero qué ocurre con la TRM. En *Conceptos*, Fodor comienza su presentación de la TRM reconociendo que dicha teoría es más bien “una vaga confederación de tesis a la que le falta una formulación canónica”¹². Sin pretensiones de subsanar esta falta, Fodor caracteriza a la TRM como la conjunción de las siguientes te-

¹² Fodor (1998), p. 24. La falta de una presentación canónica de la TRM y de sus relaciones con la TCM y la HLDP es una de las principales dificultades con las que nos encontramos aquí. Hay, por ejemplo, una notable diferencia en este respecto entre las presentaciones de Fodor (1998) y (1987), y a la vez entre éstas y exposiciones como la de Loewer y Rey (1991). A diferencia de estos autores, en este trabajo he optado por tratar a la TCM y a la TRM como tesis bien diferenciadas, ateniéndome a las definiciones para cada una de ellas que Fodor presenta en los textos citados en cada caso.

sis: (1) la explicación psicológica es típicamente nómica y completamente intencional, (2) las representaciones mentales son las portadoras primitivas de contenido intencional, (3) el pensamiento es computación, (4) el significado es información (más o menos), (5) lo que sea que distinga a conceptos coextensivos debe estar en la cabeza. Dado que para nuestros fines interesa la manera en que Fodor defiende la TRM más que la TRM en sí misma, las dos últimas tesis no son relevantes y nos concentraremos en la tercera: "el pensamiento es computación".

Ahora bien, cuando la TRM afirma que el pensamiento es computación, ¿a qué se refiere con "pensamiento"? (en otros términos, ¿qué partes de la mente deben ser tratadas como computacionales?). Una primera posibilidad es que por "pensamiento" se entienda sólo lo que ordinariamente llamamos de esa manera, i. e. la capacidad para resolver problemas complejos, la fijación de creencias, el razonamiento analógico, etc. Si este fuera el caso, la TRM estaría afirmando que los procesos mentales *centrales* son computacionales (lo cual, podrá advertirse ya, no es del todo coherente con el pesimismo de Fodor acerca de los alcances de la TCM). Una segunda posibilidad es que cuando se dice que "el pensamiento es computación", por "pensamiento" se entienda la vida mental en general, incluyendo no sólo los procesos que hemos llamado "centrales" sino también los periféricos: la percepción y los sistemas motores. Finalmente, una tercera posibilidad es que "pensamiento" se esté usando, de manera poco corriente, para hacer referencia *sólo* a los procesos periféricos. Si bien esta tercera opción es, según creo, la más consistente con los puntos de vista reconstruidos en la sección 1, claramente es inadecuada como una interpretación de lo que Fodor está afirmando aquí. Sólo las primeras dos opciones tienen alguna plausibilidad en este sentido, lo cual resulta evidente cuando se consideran las tesis (1) y (2) de la TRM. La manera en que Fodor elucida estas dos tesis nos permite ver con claridad cuáles son los estados y procesos acerca de los cuales versa la TRM. La tesis (1) implica que "los ejemplos paradigmáticos de leyes [psicológicas] son leyes acerca de las relaciones causales entre las creencias, deseos y acciones". Es decir, estados mentales propios de los sistemas centrales. Además, en la presentación de la tesis (2) Fodor introduce lo que ha sido siempre el núcleo de la TRM¹³: la idea según

¹³ Cf. p. ej. Fodor (1987), p. 43.

la cual para cada creencia (o deseo, intención, etc.) de que p por parte de un sujeto S , es necesario postular una representación mental RM con la que S guarda una determinada relación funcional, y tal que RM significa que p . Nuevamente, resulta claro que para Fodor la TRM es, principalmente, una teoría acerca de las actitudes proposicionales y, en general, acerca de los estados y procesos propios de los sistemas centrales. Lo que la TRM afirma en su tesis (3), por tanto, es que *los estados y procesos propios de los sistemas centrales son computacionales*. Como veremos luego, este resultado entra en conflicto con el pesimismo de Fodor acerca de los alcances TCM y su consecuente propuesta metodológica según la cual la ciencia cognitiva debe concentrarse en el estudio de los sistemas modulares, por ser éstos los únicos que pueden ser modelados computacionalmente.

Pasemos ahora a los argumentos con los que Fodor fundamenta la HLDP. Esta hipótesis suele ser presentada como una tesis adicional a la TRM; es decir, como una tesis que tiene sentido en el contexto de la TRM pero que no es estrictamente una consecuencia con la que un partidario de la TRM se halle necesariamente comprometido¹⁴. Según la HLDP, las representaciones mentales tienen formato lingüístico y constituyen un lenguaje con sintaxis y semántica combinatoria. De los distintos argumentos que Fodor ha presentado en favor de esta hipótesis, hay dos que resultan complementarios y que son de especial interés para nuestros fines: el *argumento de los procesos mentales* y el *argumento de productividad y sistematicidad*¹⁵. El argumento de los procesos mentales puede resumirse de la siguiente manera: nuestros compromisos ontológicos deberían ser los asumidos por las mejores teorías científicas disponibles. Las mejores teorías psicológicas disponibles son teorías computacionales, i. e. teorías que suponen que los procesos mentales pueden ser descritos en términos de computaciones sobre símbolos. Estas teorías suponen la existencia de un medio simbólico con las características del LDP sobre el cual se realizan las computaciones. Por lo tanto debemos aceptar la existencia de un medio simbólico como el LDP. Este argumento puede hallarse en casi toda la obra de Fodor, pero tiene un papel central en *El lenguaje del pensamiento* (Fodor, (1975)), buena parte del cual está destinado a mostrar de qué manera y en qué me-

¹⁴ Cf. p. ej. el apéndice de Fodor (1987); Aydede (1998).

¹⁵ Cf. el apéndice de Fodor (1987), donde estos dos argumentos se presentan de esta manera, junto con un tercero, de carácter metodológico.

dida algo así como un LDP está supuesto por distintas teorías psicológicas¹⁶.

Ahora bien, frente a este argumento que invoca el carácter computacional del pensamiento, surge nuevamente la pregunta que ya nos hicimos con respecto a la TRM: ¿qué partes de la mente deben ser entendidas como computacionales? En otros términos, ¿cuáles son los sistemas computacionales que, según el argumento de Fodor, requieren de un LDP como medio simbólico sobre el cual computar? De nuevo, una primera respuesta apunta a los sistemas centrales exclusivamente. Según esta respuesta, el LDP es el medio simbólico sobre el cual se lleva a cabo el Pensamiento con mayúscula, i. e. resolución de problemas, el razonamiento abductivo y, en general, todos los procesos que involucran actitudes proposicionales (nótese que, en este caso, la defensa de la HLDP resulta poco coherente con el pesimismo de Fodor acerca de la posibilidad de modelar computacionalmente los procesos centrales). Una segunda posibilidad es que el dominio del LDP se extienda *también* a los sistemas periféricos, mientras que una tercera consiste en que se *restrinja* a estos últimos. Interpretado según esta tercera posibilidad, el *argumento de los procesos* sería el siguiente: las mejores teorías psicológicas *acerca de sistemas periféricos* (i. e. las mejores teorías acerca de la visión, acerca de la comprensión del lenguaje, etc.) tratan a dichos sistemas como computacionales. Y dado que la computación requiere de un medio simbólico con sintaxis y semántica combinatoria, debemos aceptar la existencia de un “LDP” (o “mentalés”, como sería más apropiado llamarlo en este caso) con el cual operan las partes periféricas de la mente. Algunos textos de Fodor, parcialmente considerados, otorgan plausibilidad a esta interpretación: la presentación del argumento en Fodor (1987)¹⁷, por ejemplo, sólo hace referencia a teorías psicolingüísticas y teorías acerca de la percepción. Sin embargo, hay al menos dos razones por las cuales esta tercera interpretación debe ser descartada. En primer lugar, el arco de teorías consideradas en Fodor (1975) es considerablemente más amplio que el de Fodor (1987) e incluye teorías acerca de deci-

¹⁶ Cf. especialmente pp. 47-48, donde Fodor presenta las líneas generales de su argumento: “Lo que trataré de demostrar es que, prescindiendo de las propias suposiciones sobre los *detalles* de las teorías psicológicas de la cognición, su estructura general presupone procesos computacionales subyacentes y un sistema representacional en el que se realizan tales procesos”.

¹⁷ P. 203 y ss.

sión racional, aprendizaje de conceptos, etc.¹⁸. Si son éstas las teorías computacionales que requieren del LDP, entonces la tercera interpretación está descartada. En segundo lugar, los restantes argumentos de Fodor a favor de la HLDP dejan en claro que el LDP no es un medio simbólico con el que operen los sistemas periféricos exclusivamente. El *argumento de la sistematicidad y productividad*, por ejemplo, está basado en dos propiedades que tienen las creencias, deseos y otras actitudes proposicionales. Para nuestros fines, no interesa el argumento en sí mismo. Sólo me interesa notar el hecho de que dicho argumento se aplica fundamentalmente a *actitudes proposicionales*, es decir a estados *proprios de los sistemas centrales*. Por estas razones, la idea de que el LDP sea el código interno utilizado por los sistemas periféricos resulta inadecuada como interpretación de la posición de Fodor. Lo que sin duda Fodor afirma es que el LDP es el código interno de los sistemas centrales (y tal vez sea *también* el de los sistemas modulares, si resulta correcta la segunda de las posibilidades antes mencionadas)¹⁹.

¹⁸ “[Hay] líneas de argumentación que parecen conducir, con cierto aire de inevitabilidad, a postular un lenguaje del pensamiento como precondition para todo tipo de teorías dentro de la psicología cognitiva. Mi punto de vista es que no es sólo la acción deliberada, sino también el aprendizaje y la percepción, lo que se debe interpretar como basado en procesos computacionales; y, una vez más, no hay computación sin representación”. Fodor (1975), p. 54. Énfasis mío.

¹⁹ Vale la pena notar que algunos lectores de Fodor se han planteado de manera explícita este problema. M. Aydede, por ejemplo, afirma taxativamente que la HLDP es una hipótesis acerca de la naturaleza de las *actitudes proposicionales* y que, si bien es razonable pensar que los sistemas periféricos hacen uso de un LDP propio (dado su carácter computacional), ésta es una cuestión empírica abierta. Cf. p. ej. Aydede (1998): “HLDP es una hipótesis acerca de la naturaleza del pensamiento y de los procesos mentales con contenido proposicional. Como tal, puede o no ser aplicada a otros aspectos de la vida mental (...). Sin duda, muchos teóricos del LDP tienen puntos de vista acerca de esos aspectos de la vida mental según los cuales también ellos podrían, al parecer, ser explicados por algo similar al LDP. Fodor, por ejemplo, parece creer que muchos sistemas modulares de entradas tienen su propio LDP en la medida en que pueden ser explicados en términos representacionales y computacionales. De hecho, muchos modelos psicológicos contemporáneos tratan los sistemas perceptuales justamente de esa manera (...). Si los procesos perceptuales deben o no ser tratados dentro del marco de una HLDP completa es una cuestión empírica abierta. Podría ser que la respuesta a fuera positiva. En ese caso, podría haber más de un LDP realizado en diferentes subsistemas o mecanismos de la mente-cerebro. Por tanto, HLDP no está comprometida con que haya un único sistema representacional realizado en el cerebro (...).” Énfasis agregado.

En resumen, hemos mostrado hasta aquí que tanto la TRM como la defensa de la HLDP suponen la idea de que los sistemas centrales son computacionales. Como podrá advertirse, este resultado es difícilmente conciliable con los puntos de vista de Fodor que hemos presentado en la sección 1. El siguiente esquema resume el estado de la cuestión, presentando dos conjuntos de afirmaciones en conflicto cuyo fundamento ha sido presentado en las secciones 1 y 2 respectivamente:

Sección 1: Fodor (1983), (2000)	Sección 2: Fodor (1975), (1987), (1998)
1. La TCM (basada en la noción clásica de computación) no puede modelar el funcionamiento de los sistemas centrales, dado que dichos sistemas son quineanos e isotrópicos.	4. La TRM y la HLDP son teorías acerca de los estados y procesos mentales propios de los sistemas centrales (y tal vez también de los modulares).
2. El alcance de la TCM está acotado a las partes modulares de la mente.	5. La TRM y la HLDP suponen o afirman el carácter computacional de los estados y procesos acerca de los cuales ellas constituyen una teoría.
3. La ciencia cognitiva debe concentrarse en el estudio de las partes modulares de la mente, hasta tanto no aparezcan las ideas clave que nos permitan comprender el funcionamiento de las partes no modulares.	6. Los procesos y estados mentales propios de los sistemas centrales pueden ser descritos en términos computacionales (esto es una consecuencia de haber aceptado 4 y 5).

Como puede verse, si la tesis (6) es verdadera, las tesis (1) y (2) no pueden serlo, con lo cual pierde sentido la recomendación metodológica hecha en (3). Esta situación puede invitar a la búsqueda de una interpretación más caritativa de las afirmaciones de Fodor que la propuesta hasta aquí. La manera más simple de salvar la consistencia de sus puntos de vista consiste en rechazar alguna de las afirmaciones del cuadro. Sin embargo, no me parece que esto sea posible: por un lado, las tesis (1), (2) y (3) constituyen el mensaje central de Fodor (2000), y su afirmación es reiterada a lo largo del libro. Por otro lado, si mi argumentación en esta sección ha sido correcta, tampoco es posible negar que Fodor se comprometa con (4), (5) y (6).

Una segunda manera de salvar la consistencia es denunciar un equívoco en las afirmaciones (1) y (6). Según esta sugerencia, mientras que (1) alude a la TCM y a la noción *clásica* de computación, (6) hace referencia, de manera mucho más vaga, a "computa-

ciones” y “teorías computacionales”. Esto deja abierta la posibilidad de que haya teorías computacionales distintas a la TCM, teorías computacionales que no supongan la noción *clásica* de computación sino alguna noción alternativa. Dada esta posibilidad, (6) no afirma *exactamente* lo mismo que (1) niega y no hay inconsistencia alguna. En mi opinión, tampoco esta segunda estrategia resulta eficaz. Si bien es cierto que el pesimismo de Fodor (2000) está puntualmente dirigido a la noción *clásica* de computación (i. e., a la noción de computación como transformación sintáctica de símbolos) y si bien es cierto que deja abierta (o al menos no cierra explícitamente) la posibilidad de elaborar una noción alternativa, no hay indicios en toda su obra de cuál pueda ser esa noción. Y, de todas maneras, la noción de computación a la que él mismo apela en su defensa de la TRM y la HLDP no es otra que la clásica. Por dar sólo un ejemplo, cuando Fodor (1998) afirma que según la TRM “el pensamiento es computación”, se remonta una vez más al legado de Turing, dejando en claro que la noción de computación que está suponiendo es la noción clásica²⁰. De manera que la propuesta de elaborar una noción alternativa no constituye para Fodor una salida al conflicto entre (1) y (6), por mucho que sea su interés intrínseco.

Podemos concluir, entonces, que hay una tensión no resuelta entre las dos líneas de investigación presentes en la obra de Fodor: la línea pesimista de Fodor (1983) y (2000) entra en conflicto con los argumentos clásicos a favor de la TRM y la HLDP presentes en Fodor (1975), (1987) y (1998). Y las estrategias más visibles a las que un fodoriano podría apelar para resolver el conflicto no son, según hemos visto, del todo satisfactorias. Desde mi punto de vista, Fodor tiene pendiente una explicación más detallada de la manera en que sus distintas contribuciones a la filosofía y a las ciencias cognitivas conforman una imagen única y coherente del funcionamiento de la mente. Su nuevo libro confirma la necesidad de dicha explicación, sobre todo para aquellos lectores que, mayormente provenientes del ámbito filosófico, encontrarán especialmente sorprendentes sus afirmaciones acerca de la incapacidad de la TCM para modelar los fenómenos más característicos de la inteligencia humana.

²⁰ Cf. Fodor (1998), p. 28.

BIBLIOGRAFIA

- Aydede, M. (1998), "Language of Thought Hypothesis", *Stanford Encyclopedia of Philosophy*: <http://plato.stanford.edu/entries/language-thought/>
- Fodor, J. A. (1975), *The Language of Thought*. Harvard University Press. Referencias a la trad. cast.: *El lenguaje del pensamiento*, Madrid, Alianza, 1984.
- (1983), *The Modularity of Mind*, Cambridge Mass. MIT Press. Referencias a la trad. cast.: *La modularidad de la mente*, Madrid, Morata, 1986.
- (1987), *Psychosemantics: the Problem of Meaning in the Philosophy of Mind*, Cambridge Mass. MIT Press. Referencias a la trad. cast.: *Psicosemántica*, Madrid, Tecnos, 1994.
- (1990), *A Theory of Content*, Cambridge Mass, MIT Press.
- (1994), *The Elm and the Expert. Mentalese and its Semantics*, Cambridge Mass, MIT Press.
- (1998), *Concepts: Where Cognitive Science Went Wrong*, N. Y., Oxford University Press. Referencias a la trad. cast.: *Conceptos: donde la ciencia cognitiva se equivocó*, Barcelona, Gedisa, 1999.
- (2000), *The Mind Doesn't Work that Way*, Cambridge, MIT Press.
- Loewer B. y Rey G. (1991), "Editor's Introduction" en *Meaning in Mind: Fodor and his Critics*, Oxford, Blackwell.
- Pinker, S. (1997), *How the Mind Works*, New York, Norton.
- Plotkin (1997), *Evolution in Mind*, London, Alan Lane.