

**Tesis de Licenciatura**

---

**ATRIBUCIONES INTENCIONALES  
Y COMPETENCIA LINGÜÍSTICA  
UN ENFOQUE EVOLUCIONISTA**

---

**Nicolás Venturelli**

**Dra. Carolina Scotto**  
–directora de tesis–

**Universidad Nacional de Córdoba**  
**Facultad de Filosofía y Humanidades**  
**Escuela de Filosofía**

**Mayo 2006**

## Introducción

Este trabajo trata sobre perspectivas; trata sobre orientaciones, rumbos, enfoques. Asumo sin argumento que ningún teórico y que ninguna investigación, de los tipos que sean, “parten de cero”: ningún teórico y ninguna investigación son, por así decir, inocentes. Esto es especialmente así para el caso de los investigadores de “la mente”, un concepto tan ubicuo e indefinido como puede haber, un concepto que fácilmente admite interpretaciones diferentes e incompatibles, y cuyos alcances y límites son, cuanto menos, borrosos: un concepto, en fin, que confiere al que quiera abordarlo, del modo que sea, una libertad considerable. Por otra parte, incluso para los objetivos más declarada y, sin duda, en algunos casos, laudablemente “científicos” (admitiendo toda la laxitud del asunto), el concepto en cuestión es importante, fundamental, imprescindible.

Esto no es aceptado unívocamente, ni mucho menos. En muchos círculos académicos, todavía hay quienes creen (y argumentan) que todo lo que se haya dicho o pueda llegar a decirse sobre el comportamiento, en el más amplio sentido, de los seres vivos puede prescindir de este resbaladizo concepto (y, claro está, de una hueste de otros conceptos, que caen bajo su jurisdicción). Pero aceptar aquello, y negar esto, es, sin duda, otro de los supuestos centrales sobre los que este trabajo se asienta. Para ya entrar en la jerga filosófica, podemos decir que el alcance explicativo (y, por qué no, expresivo) del conjunto de conceptos llamados “mentalistas” no puede, ni pudo ni podrá, reducirse a un vocabulario o una terminología “fiscalista”, que refiera, esto es, estrictamente a entidades, propiedades y / o procesos físicos.

A la luz de lo dicho, esto es, por una parte, ante la *dificultad* de hablar sobre la mente y, por otra parte, ante la *necesidad* de hacerlo, el análisis filosófico entra de lleno en el campo de acción. Para decirlo en una oración, a grandes rasgos, el objeto de estudio de este trabajo son justamente algunos de los cuestionamientos filosóficos acerca de cómo el concepto de mente pueda integrarse, o no, a una perspectiva teórica que sea científicamente viable y, sobre todo, útil para las demás disciplinas relacionadas, de un modo u otro, con el estudio de la mente (estoy pensando especialmente en la psicología cognitiva en general, la lingüística, la etología cognitiva,

la psicología del desarrollo y la investigación en Inteligencia Artificial, entre otras). Esta tarea será enfocada desde un punto de vista eminentemente filosófico, en la discusión de una cantidad de supuestos teóricos y metodológicos que influyen, en general y en particular, en un conjunto de problemas relacionados y comunes a dos propuestas filosóficas sustantivas de inmenso peso en la filosofía contemporánea de la mente (en este sentido, no me centro especialmente en los problemas propios de, y las importantes ramificaciones que dichas propuestas tienen en, otras disciplinas).

Las propuestas filosóficas en las que me concentraré son la llamada “Teoría de los Sistemas Intencionales”, de Daniel Dennett, y la llamada “Teoría Representacional de la Mente”, de Jerry Fodor. Estas son específicamente *teorías filosóficas de la intencionalidad*. Mi interés está puesto particularmente en el análisis e interpretación de la primera de estas posturas. En la medida que esto es así, el estudio de la teoría de Fodor será, hasta cierto grado, instrumental y, de acuerdo con esto, haré hincapié especialmente en su contraste con la teoría dennettiana. De hecho, una de las razones por las que elijo a Fodor para esto tiene que ver con el modo en que difiere de la teoría de Dennett. Lo que quiero decir es lo siguiente: las teorías en cuestión no sólo son, individualmente, y sin lugar a duda, dos de las posiciones más influyentes en el campo, dentro y fuera de la filosofía de la mente, sino que también, de modo muy claro, puede visualizarse entre ambas una situación particular, que llamó mi atención desde el momento en que emprendí su estudio.

El intenso debate explícito que Dennett y Fodor han estado llevando adelante se remonta casi hasta los primeros trabajos de ambos autores y perduró a través de un gran número de publicaciones a lo largo de los años. Aún así, a pesar de algunas revisiones internas menores, ninguno de ambos filósofos ha cedido terreno al contendiente en ninguno de los frentes bajo crítica, hasta el punto en que cabe preguntarse si de hecho la contienda es fructífera o en realidad termina asemejándose a un improductivo diálogo filosófico de sordos. Muy burdamente, dado el modo en que desarrollan sus respectivas teorías de la intencionalidad, daría la impresión de que no están dando cuenta del mismo fenómeno. En otras palabras, puede atribuirse una fuerte incompatibilidad “en bloque”, que excede ampliamente las discusiones particulares, entre ambas teorías.

Por otra parte, hay que notar cómo el desarrollo teórico dennettiano se apoya en muchísimas ocasiones, para diversas temáticas y en diversos contextos, de modo explícito e implícito, en el contraste con, y la crítica a, las principales tesis fodorianas. Indudablemente, la teoría representacional de la mente ha sido, en este sentido acotado,

una gran fuente de inspiración para Dennett y su importancia debe reconocerse en la trayectoria de las ideas de este último. Fodor puede considerarse, sin lugar a dudas, el “anti-héroe” dennettiano por excelencia.

Ahora bien, como mencioné antes, tanto la teoría de los sistemas intencionales como la teoría representacional de la mente pueden considerarse, y así también son (a grandes rasgos) entendidas por sus autores, teorías filosóficas de la intencionalidad. La noción de intencionalidad es históricamente asociada con el trabajo de Franz Brentano. Como es bien sabido, Brentano consideraba la intencionalidad como el rasgo distintivo que diferencia los fenómenos mentales de todo otro fenómeno. Aún así, la intencionalidad es en general caracterizada más ampliamente como un “ser acerca de”, como, digamos, una direccionalidad o referencialidad (en este sentido, también se la llama a veces “acerquidad”) y consiste, fundamentalmente, en la capacidad, propia de cierta clase de entidades, para referir a, o ser acerca de, otra entidad (o conjunto de entidades) o alguna de sus propiedades (o conjunto de ellas). Sin incorporar el bagaje teórico brentaniano, me interesa aquí la noción de intencionalidad en tanto que acotada al marco de los fenómenos mentales y, en este sentido, a la cuestión de su *contenido intencional*.

En este marco, una de las discusiones contemporáneas que ha tenido una enorme repercusión en la literatura, y que en mi opinión es central para los planteos dennettiano y fodoriano, es la que se ocupa de la naturaleza de la así llamada “psicología popular” o “psicología de sentido común”. El término “psicología popular” (*folk psychology*) ha sido acuñado por el mismo Dennett (1981b) y ha entrado a formar parte del vocabulario técnico estándar de los filósofos de la mente. Se entiende por tal, en términos generales, el conjunto de conceptos y generalizaciones mediante los cuales las personas cotidianamente atribuimos vida mental a otros organismos (nuestros congéneres así como otros agentes) con el objeto de entender, predecir y / o explicar su comportamiento; esta atribución implica de modo central la apelación a estados mentales como creencias y deseos (pero también intenciones, dudas, temores, esperanzas, etcétera). El discurso psicológico popular ha generado discusión no sólo por el interés que su estudio tiene por sí mismo, sino en tanto que ha probado ser uno de los enfoques más fructíferos para abordar las cuestiones acerca de la intencionalidad de la mente.

El modo en que las teorías de la intencionalidad dennettiana y fodoriana abordan las cuestiones relativas a la psicología popular resultará ser finalmente el punto focal de

disputa entre ambas, de acuerdo a la particular interpretación que haré de las teorías. El objetivo puntual de mi trabajo es justamente aportar un determinado enfoque desde el cual dichas teorías pueden ser interpretadas. A la luz del enfoque propuesto, podrá luego darse cuenta de la mencionada incompatibilidad en bloque de las teorías, pero fundamentalmente se configurarán dos modelos filosóficos generales para el abordaje de cualquier temática vinculada a la mente. El enfoque en cuestión consiste específicamente en la medida en que una teoría filosófica de la mente mantenga, o no, alguna relación con la teoría de la evolución por selección natural<sup>1</sup>, propuesta por Charles Darwin (1859).

La tesis vertebradora de este trabajo es la siguiente: hay un marco metodológico y teórico, básico y general, que puede rastrearse en la teoría de los sistemas intencionales de Dennett, que se asienta en el peso filosófico que ésta atribuye a la teoría evolucionista darwiniana. En tanto que este marco puede interpretarse como teniendo un rol substancial en la interna de la disputa entre las teorías dennettiana y fodoriana, habremos obtenido una herramienta meta-teórica que, sostengo, es imprescindible para dirimir dicha disputa, especialmente considerando el altísimo grado en que las teorías son incompatibles. En acuerdo con esto, la teoría dennettiana es, en un sentido importante, un intento de mostrar que el abordaje desde el mencionado marco de toda cuestión relativa al carácter intencional de la mente, es no sólo beneficioso sino ineludible. Denomino el marco en cuestión “enfoque evolucionista acerca del contenido mental” y lo opongo a lo que llamo “enfoque anti-evolucionista acerca del contenido mental”, que atribuyo a la teoría representacional de la mente.

Para la defensa de mi tesis general, me apoyo en una distinción crucial, que a mi entender merece particular atención y que ha sido generalmente pasada por alto o interpretada de modo parcial. Me refiero a la distinción técnica propuesta por Dennett en 1977<sup>2</sup> entre “creencias” y “opiniones”, distinción que será central para relevar los diversos aspectos del enfoque evolucionista acerca del contenido mental y, especialmente, el modo en que éstos cumplen un rol filosófico en la disputa con la teoría fodoriana. Puntualmente, la distinción sienta las bases para remover un supuesto de homogeneidad que en general impregna dicha teoría. Este supuesto es implícitamente modelado sobre la base del reporte lingüístico en primera persona de pensamientos con

---

<sup>1</sup> Aunque sea considerada una tarea fundamental y necesaria, no me ocupo especialmente de justificar una determinada interpretación de la teoría darwinista (y, en particular, de dirimir las cuestiones que enfrentan los neo-darwinistas o adaptacionistas a sus críticos).

<sup>2</sup> En Dennett (1978a).

contenido proposicional preciso y complejo, propio de los hablantes competentes de un lenguaje articulado.

La distinción entre creencias y opiniones constituye una importantísima herramienta teórica que no solamente no entra en conflicto con el enfoque evolucionista dennettiano, sino que justamente vuelve inteligible el tipo de fenómeno que aquel pretende explicar (o, más precisamente, cree posible y justificable explicar), haciendo por otra parte manifiesto el abanico de supuestos cuestionables y confusiones que llevan a los defensores de un enfoque anti-evolucionista a configurar su objeto de estudio. A la vez, dada la fundamental importancia que, como sostengo, el enfoque evolucionista cobra para una adecuada interpretación de la postura dennettiana, mantengo que la distinción entre creencias y opiniones constituye uno de los principales ejes teóricos de dicha postura. Es importante aclarar aquí que mi interés en la distinción reside tanto en enfocar su análisis para remedar el grave sub-desarrollo de la misma en la literatura (incluso por parte del mismo Dennett), como en hacerlo a los fines de desplegar el alcance del evolucionismo dennettiano.

Ahora bien, la relevancia de los argumentos defendidos en este trabajo excede ampliamente el contexto acotado de la disputa entre Dennett y Fodor. El nivel de análisis al que mayormente apunta mi argumentación es aquel donde la teorización sustantiva se asienta y se encuadra. En particular, por un lado, la crítica al enfoque anti-evolucionista acerca del contenido mental conllevará un ataque a un marco metodológico supuesto en buena parte de la discusión contemporánea en filosofía de la mente: esto es, el análisis filosófico en términos de las llamadas “actitudes proposicionales”. Por otro lado, la defensa del enfoque evolucionista delimitará una perspectiva metodológica relevante tanto para el estudio filosófico de la mente en general como para el conjunto de disciplinas aplicadas vinculadas al estudio de la mente.

El presente trabajo se subdivide en cuatro capítulos (y una conclusión). En los capítulos 1 y 2 presento las teorías dennettiana y fodoriana de la intencionalidad, con un énfasis particular en el relevamiento de sus aspectos metodológicos y pre-argumentativos, lo que incluye especialmente sus supuestos generales, alcances, límites, herramientas conceptuales y problemáticas abarcadas, a los fines de justificar su inserción respectiva en los enfoques evolucionista y anti-evolucionista acerca del contenido mental.

En el capítulo 3 me ocupo de desarrollar y analizar a fondo las nociones técnicas de “creencia” y de “opinión”, tratando especialmente de dirimir los problemas acerca de su interpretación. En el marco de esta tarea, intento hacer frente a las acusaciones de anti-gradualismo a la distinción, tratando a la vez de compatibilizarla con el enfoque evolucionista dennettiano. Finalmente, exploro los alcances de la hipótesis según la cual la distinción es un elemento nuclear de la teoría de los sistemas intencionales, justamente en tanto que puede funcionar como herramienta teórica al servicio del enfoque evolucionista acerca del contenido mental.

El capítulo 4 consiste, en su mayor parte, en la confirmación argumentada de esta hipótesis. De acuerdo al desarrollo de la argumentación, en primer lugar, me dedico a las concepciones del discurso psicológico popular que, según sostengo, pueden atribuirse a la teoría de los sistemas intencionales y a la teoría representacional de la mente; en segundo lugar, evalúo las implicaciones y la adecuación del análisis en términos de actitudes proposicionales; en tercer lugar, me concentro en las cuestiones sustantivas acerca de la configuración de los fenómenos mentales y de la determinación de su contenido que las teorías defienden; en cuarto lugar, analizo la repercusión meta-teórica, en relación con emprendimientos teóricos no filosóficos, que puede perfilarse en las teorías estudiadas.

Quisiera hacer una aclaración final, antes de comenzar. Tal como este trabajo fue planteado, se vuelve necesaria una perspectiva de análisis “verticalista”, por así decir. Me refiero a que efectuaré un trabajo analítico en profundidad, centrado casi exclusivamente en (algunos puntos de) las obras de Dennett y de Fodor. Al tratar algunos puntos específicos desde una perspectiva de análisis particular, un estudio amplio y abarcador es simplemente prohibitivo. A la vez, en la medida que el enfrentamiento entre Dennett y Fodor, en lo que respecta a la relación que mantienen con la teoría evolucionista, es en general mantenido tácito, se hace apremiante el trabajo analítico de escarbar en sus fundamentos e implicaciones.

Manos a la obra.

---

## La Teoría de los Sistemas Intencionales

### *1. Presentación General de la Teoría*

En este apartado presentaré lo que es considerado el núcleo de la posición de Dennett acerca del contenido mental, esto es, la llamada “Teoría de los Sistemas Intencionales”; esto se realizará a los fines de relevar el marco metodológico general en el que se circunscribe Dennett para elaborar su postura, tarea que desarrollaré en los subsiguientes dos apartados.

Dennett es un autor paradigmáticamente ecléctico cuyo peculiar método filosófico (que recuerda en algunos aspectos al estilo wittgensteineano) revela un claro desinterés en la construcción de teorías filosóficas sistemáticas o incluso cuerpos teóricos unitarios y cerrados, mucho menos si éstos son puramente filosóficos. Este método lo lleva a brindar aportes teóricos más o menos dispersos y de diverso tipo, incorporando diferentes perspectivas y fuentes multi-disciplinarias<sup>3</sup>. Dada la inclinación metodológica dennettiana, se hará apremiante la tarea de reconstruir los diversos aportes teóricos de nuestro autor en un cuerpo más o menos integrado<sup>4</sup>. Aún así, puede circunscribirse el cuerpo central de su postura acerca del contenido mental en una serie de artículos, reunidos mayormente en las dos primeras colecciones, *Brainstorms* (1978a), que incluye “Intentional Systems” (1971), el primer artículo claramente inscripto en lo que de modo paradigmático se considera la teoría dennettiana del contenido mental, y *The Intentional Stance* (1987a), el texto dennettiano por excelencia sobre contenido mental<sup>5</sup>.

El eje central alrededor del cual gravitan muchas de las ideas dennettianas acerca del contenido mental es lo que en “True Believers” (1981a) el filósofo llama “estrategia intencional” o “actitud intencional” (*intentional stance*). La actitud intencional es una

---

<sup>3</sup> En el apartado 1.2 retomaré la temática del método dennettiano analizándolo como una toma de posición no ingenua y teóricamente relevante.

<sup>4</sup> Es sólo en este sentido laxo, de hecho, que puede llamarse “teoría” a la propuesta dennettiana.

<sup>5</sup> Además de los textos mencionados, trabajaré sobre *Kinds of Minds* (1996a), algunos artículos pertinentes de la colección titulada *Brainchildren* (1998a), y, en menor medida, en el texto seminal (y quizás todavía no del todo maduro) de Dennett, *Content and Consciousness* (1969).

entres otras actitudes o estrategias que los seres humanos adoptamos en nuestro esfuerzo para comprender el mundo que nos rodea. Para este fin, podemos adoptar además la actitud física, que dirige su atención a la constitución física del objeto estudiado y a sus modificaciones de acuerdo con las leyes de la física, y la actitud de diseño, que considera al objeto como diseñado e inquiera sobre su comportamiento de acuerdo a los fines para el cumplimiento de los cuales el mismo fue diseñado. Cada una de estas actitudes tiene características distintivas y, concordemente, su adopción comporta ventajas y desventajas particulares.

La actitud intencional tiene un alcance mucho más restringido que la actitud de diseño (que puede adoptarse frente a cualquier objeto cuyo funcionamiento exhiba un diseño) y que la actitud física (que puede adoptarse frente a cualquier objeto físico, independientemente de su comportamiento). Aún así, la actitud intencional puede considerarse un subtipo de la actitud de diseño y la diferencia en los alcances de ambas estrategias depende exclusivamente de las características particulares de la puesta en práctica de la actitud intencional y del éxito predictivo y explicativo que ésta confiere sobre la adopción de la actitud de diseño.

La actitud intencional consiste básicamente en tratar al objeto cuyo comportamiento estamos tratando de predecir y / o explicar como un agente racional, esto es, un agente que exhibe creencias, deseos y demás fenómenos mentales intencionales. Hecho esto, se procede a deducir qué creencias dicho agente tendría que tener dada su posición en el mundo, sus objetivos y necesidades, y, mediante las mismas consideraciones, qué deseos tendría que tener. La asunción de racionalidad implicará luego que el agente se comportará de acuerdo a sus creencias para la satisfacción de sus deseos: esta asunción es el elemento crucial distintivo de la estrategia.

La tesis central de Dennett en este punto es que el contenido (esto es, el mismo carácter intencional) de dichos fenómenos mentales es fijado por medio de la adopción de la estrategia intencional: tal estrategia es de hecho el único medio para la fijación del contenido mental, hecho que la convertiría en un mecanismo absolutamente imprescindible para el abordaje de todos los problemas relativos a lo que Dennett (1981b) denominó “psicología popular” (*folk psychology*) o psicología de sentido común. En este sentido, la creencia, el deseo y demás son fenómenos que solamente pueden discernirse por medio de la adopción de tal estrategia predictiva: cualquier sistema cuyo comportamiento pueda ser de este modo pronosticado por esta estrategia es, en su más completo sentido, un “sistema intencional”.

La característica que otorga la ventaja crucial de la actitud intencional por sobre de las demás actitudes es su *simplicidad*. La cantidad de factores que pueden entrar en juego al adoptar la actitud física (en la determinación de múltiples relaciones causales de acuerdo a las leyes de la física) y la actitud de diseño (en la determinación de procesos de funcionamiento de acuerdo a la estipulación del diseño) es reemplazada por la atribución directa de creencias y deseos particulares, que condensan en una única postulación aspectos puntuales de una realidad que, desinterpretada, puede ser muy compleja y heterogénea. Esta ventaja puede explicarse de otra forma apelando a la distinción, propuesta primeramente por Dennett<sup>6</sup>, entre un nivel *personal* y un nivel *sub-personal* de explicación.

La práctica atributiva que se pone en juego mediante la estrategia intencional refiere al comportamiento de los organismos en tanto que son tomados como sistemas totales, por lo tanto embarcados en actividades de nivel personal: de acuerdo con esto, la atribución de contenido mental en ningún caso necesita bajar al nivel sub-personal, esto es, aquel referido a los mecanismos internos de implementación de los estados atribuidos<sup>7</sup>. Resulta obvio que la abstención de apelar a entidades o propiedades cuya postulación se inserta en la probablemente complicadísima estructura cognitiva interna del sistema confiere economía explicativa, más aún cuando, en lugar de esta postulación, la explicación se rige por la apelación a elementos altamente condensados y no rígidos como son las creencias y los deseos<sup>8</sup>. Esto a su vez implica que la actitud intencional está como tal imposibilitada para la predicción precisa de acciones; esta imprecisión es, sin embargo, justamente la fuente del alto poder predictivo de la estrategia, que en casos complejos puede obviar los detalles de la situación simplificando la atribución.

Un punto central relacionado con lo anterior (y que será retomado más adelante) es el hecho de que la atribución de creencias (deseos, etcétera) depende de la constatación de lo que podría llamarse la “situación ambiental” del organismo, esto es, su inmersión en el mundo dictada por la experiencia adquirida a nivel personal y también, como veremos en el apartado 1.3, por el transcurso evolutivo que desemboca en él. En este sentido, no sólo la interpretación toma al organismo como un sistema

---

<sup>6</sup> Dennett (1969, pp. 125 y siguientes).

<sup>7</sup> El problema puntual de la relación entre ambos niveles de explicación será tratado en el apartado 4.4.

<sup>8</sup> Hay que aclarar que esta caracterización de las creencias y deseos no es, en lo más mínimo, compartida por muchos oponentes de la teoría dennettiana (especialmente, para nuestro caso, Fodor).

unitario sino que lo toma en su relación con el ambiente en el que está inserto y con el que interactúa.

Las variaciones conductuales del organismo cuyo comportamiento se quiere predecir, relacionadas con las contingencias ambientales próximas a él, son los únicos datos (empíricos, si se quiere) necesarios para poner en juego la estrategia intencional, además de algunas otras verdades básicas sobre el organismo<sup>9</sup>. La complejidad de la conexión con el mundo que el comportamiento de un organismo exhibe está por este motivo en relación directa con la semántica de los estados mentales atribuibles: en cuanto aumenta la complejidad de las interacciones organismo / mundo se reduce el espectro de los posibles modelos explicativos satisfactorios, modelos que de este modo requerirán una semántica más rica y, por lo tanto, un mayor especificación del contenido mental atribuido.

La imagen que este modo de explicar la atribución de contenidos mentales nos deja de la psicología de sentido común es, con palabras de Dennett, la de un “cálculo racionalista de interpretación y predicción”<sup>10</sup>. De acuerdo con esto, la psicología popular es idealizada, abstracta e instrumental: es idealizada en tanto que sus explicaciones y predicciones son el producto de una asunción de racionalidad, esto es, de consideraciones normativas; es abstracta en tanto que los estados particulares que atribuye (las creencias y deseos) no describen correspondientes estados salientes de un sistema interno causalmente responsable de la conducta (en otra terminología, creencias y deseos no son vistos como sucesos o eventos acontecidos sino, justamente, como estados)<sup>11</sup>; es instrumental en tanto que su puesta en acción es motivada por necesidades prácticas de convivencia y colaboración, esto es, de índole esencialmente social.

Otro aspecto de esto último es que todos nosotros aprendemos a usar la psicología popular como una técnica social, esto es, como una habilidad o una destreza (*craft*), y no explícitamente como una teoría, con la identificación de entidades y procesos fijos que obedecen a leyes o regularidades establecidas<sup>12</sup>. Este aspecto netamente pragmático del conjunto de productos de la actitud intencional viene a coincidir con la imprecisión típica que antes señalábamos respecto de su funcionamiento. Los términos de la psicología de sentido común no pasan, ni necesitan

---

<sup>9</sup> Puede verse en este punto un estrecho parecido con la situación hipotética del traductor radical quineano frente a una comunidad hablante totalmente desconocida.

<sup>10</sup> Dennett (1987a, p. 55).

<sup>11</sup> Retomo este punto en la primera parte del apartado 4.3.

<sup>12</sup> Este será un punto fundamental para mi discusión de los enfoques evolucionista y anti-evolucionista acerca del contenido mental: será desarrollado en pp. 65-66 y especialmente en el apartado 4.1.

pasar, por un proceso analítico de “limpieza” conceptual a los fines de preservar su eficacia. Más aún, esta misma eficacia es un *resultado* de la impureza de los términos mentalistas y de sus operaciones.

Un blanco importante de las críticas a la propuesta dennettiana refiere a la medida en que su idea de una actitud intencional impide dirimir la cuestión (especialmente delicada para el crítico filosóficamente entrenado) acerca de un realismo o un anti-realismo (o lo que puede también llamarse “interpretacionismo”) respecto de la creencia y de los demás estados intencionales, ambas posturas tradicionalmente consideradas como diametralmente opuestas y excluyentes. En “Real Patterns” (1991c) Dennett denomina su propia posición “realismo moderado” (*mild realism*), a la vez dejando en claro su antipatía hacia el uso de rótulos (los famosos “ismos” filosóficos) que tienden a forzar las discusiones y diluir las distinciones. En acuerdo con esto, y para concluir esta presentación, trataré este problema de la “ubicación” de la teoría de los sistemas intencionales sólo en la medida en que dicha cuestión esclarezca aspectos del planteo general de Dennett.

Las aludidas críticas refieren al núcleo de la propuesta dennettiana, la tesis de que “mientras la creencia es un fenómeno perfectamente objetivo [...], puede ser discernido solamente desde el punto de vista de alguien que adopta cierta estrategia predictiva, y cuya existencia puede ser confirmada sólo por una evaluación del éxito de esa estrategia...”<sup>13</sup>. La actitud intencional permite la construcción de modelos del comportamiento a la base de los cuales pueden producirse las generalizaciones y las predicciones del lenguaje mentalista; estos modelos, en tanto identifican patrones que reúnen factores comportamentales, ambientales y disposicionales observables, son objetivos: esto es, describen un estado de cosas que efectivamente se da en el mundo y lo hacen de modo comprobable y, por lo tanto, falseable.

El hecho de que la actitud intencional sea la única estrategia predictiva capaz de configurar esos modelos objetivos la vuelve imprescindible toda vez que se quiera explicar o predecir la conducta inteligente de un sistema. En este sentido, la actitud intencional es una estrategia a la cual, de una forma u otra, debe apelarse en cualquier emprendimiento teórico que pretenda dar cuenta de la conducta inteligente, en su más amplio sentido; esto equivale a decir que las ciencias de la conducta (especialmente una psicología cognitiva madura), con sus inapelables pretensiones de objetividad, no

---

<sup>13</sup> Dennett (1987a, p. 26).

pueden prescindir de este mecanismo. A la vez, la exclusividad de los resultados teóricos de la estrategia (esto es, el hecho de que únicamente por medio de su adopción los patrones denotados por los términos mentalistas se vuelvan salientes) los vuelve inmediatamente irreducibles a un lenguaje más básico<sup>14</sup>.

Ahora bien, hay que tener en claro que, para Dennett, no hay cosas que por sí mismas sean sistemas intencionales, sino que lo central es la posibilidad de adoptar la actitud atributiva respectiva y, fundamentalmente, su éxito predictivo. Al poner fuera de consideración el hecho de poseer o no alguna determinada configuración psicológica que efectivamente instancie el estado mental atribuido, la atribución intencional es descartada sólo cuando la estrategia de atribución fracasa, esto es, cuando sencillamente no es útil para los fines que fue invocada: no aporta poder predictivo o explicativo alguno. No hay consideraciones ulteriores para la adecuación de las atribuciones particulares de contenido mental.

A la vez, es importante recalcar que el éxito de la predicción dependerá de modo directo de cómo la construcción del modelo predictivo de hecho logre dar cuenta más o menos fielmente de lo que llamaba la “situación ambiental” del organismo. Después de todo, los modelos describen estados de cosas. Digo “más o menos fielmente” porque, como he mencionado antes, la actitud intencional está sistemáticamente imposibilitada (dado su funcionamiento) para efectuar atribuciones precisas. Esto implicará puntualmente algún grado de sub-determinación de la teoría por los datos (la tesis Duhem-Quine, 1951): dada la imprecisión característica de los términos invocados para la construcción de los modelos explicativos, en principio siempre será posible la coexistencia de dos o más modelos diferentes, sin que haya ningún hecho ulterior que pueda dirimir la cuestión.

Otra crítica relacionada con la acusación de perspectivismo radical es que pareciera que el agente al cual se adscriben los contenidos intencionales no tendría, de acuerdo al modelo dennettiano, ningún tipo de autoridad sobre sus propios estados en tanto que en él se descarta la determinación intrínseca de dichos estados en función de criterios externalistas. Este problema se ha convertido en un importante debate entre los defensores de la llamada intencionalidad intrínseca (u original) *versus* los defensores de la intencionalidad derivada. Dennett es un acérrimo defensor de la segunda posición y

---

<sup>14</sup> El tema históricamente muy recurrente de la reducción del vocabulario mentalista a términos fisicalistas no será un tema puntual de discusión en este trabajo, aunque en Dennett es especialmente central en la primera parte de *Content and Consciousness*.

su solución al problema invoca explícitamente la teoría darwinista de la evolución por selección natural. Como esta discusión no es imprescindible para la presentación del núcleo de la teoría de los sistemas intencionales a la que aquí me he dedicado, y como la relación de dicha teoría con la teoría evolucionista será tratada en detalle en el apartado 1.3, dejaré para ese momento el análisis de esta discusión. En lo que sigue intentaré relevar el importante carácter *gradualista* de la teoría dennettiana aquí presentada, en sus diversos aspectos.

## **2. Carácter Gradualista de la Teoría**

A pesar de que lo que llamo “gradualismo” no es una posición filosófica unitaria o siquiera reconocida como tal, el término me será especialmente útil para demarcar una serie de aspectos de la teoría de los sistemas intencionales de Dennett y de sus consecuencias para los problemas especiales en las discusiones filosóficas sobre contenido mental que más adelante trataré. El relevamiento de dichos aspectos será la base para la inserción de dicha teoría dentro de lo que voy a denominar “enfoque evolucionista acerca del contenido mental”, cuyas características y justificación serán el tema del próximo apartado.

Voy a defender en el presente apartado que la teoría de los sistemas intencionales puede describirse justificadamente como gradualista, y no solamente en base al análisis de sus tesis sustantivas. Con esto último quiero decir especialmente que la teoría está inserta en un planteo metodológico gradualista, lo cual incluye sus supuestos, alcances y objetivos teóricos, e incluso su estilo argumentativo general. Comenzaré relevando su carácter gradualista sustantivo para después concentrarme en el método dennettiano.

El primer punto que quiero destacar es probablemente el más relevante y está implícito en el mecanismo de la actitud intencional. Como vimos, la definición de lo que para Dennett es un “sistema intencional” (esto es, todo sistema al que pueden adscribirse predictivamente estados intencionales) implica la adopción de una estrategia particular y, por esto, implica centralmente la perspectiva externa de un intérprete. Ser un sistema intencional no depende de algo que pueda determinarse definitivamente apelando a características intrínsecas de la constitución física del organismo: depende de la adopción de una determinada perspectiva a través de la que se sistematizan una

serie de patrones con fines explicativos y descriptivos. La creencia, desde esta visión perspectivista, es antes que nada un resultado: esto es, resultado de nuestra adopción de la estrategia intencional<sup>15</sup>.

Esta inexorable dependencia del intérprete (y de su peculiar toma de posición frente al sistema interpretado) respecto de la determinación del estatus de “sistema intencional” implica la aplicabilidad de la estrategia intencional a una gama de agentes en principio (sólo en principio, claro está) ilimitada. Esto en realidad sería característico de toda posición perspectivista respecto de los estados mentales: si la determinación de éstos depende únicamente de la adopción de una perspectiva (y de su éxito predictivo / explicativo) no hay (de nuevo) en principio restricción sobre los posibles agentes hacia los cuales la perspectiva se dirige. Este punto tiene una importante derivación metodológica, que más adelante retomaré.

Como ya había señalado, el factor que acota el espectro de los sistemas intencionales es (exclusivamente) la adecuación predictiva y explicativa de la estrategia, esto es, su éxito. Este éxito depende de forma directa del modo cómo las atribuciones intencionales logran condensar, apelando a un principio idealizador de racionalidad, la de otro modo posiblemente compleja “situación ambiental” del organismo (ver p. 9), es decir, el conjunto de las interacciones del organismo con su entorno dadas las variaciones del mismo, la historia evolutiva del organismo y su experiencia personal adquirida. Dado esto, la particular configuración dennettiana del fenómeno “estado intencional” (no supuesto sino configurado *via* la actitud intencional) está en estrecha relación con un conjunto complejo de factores comportamentales y ambientales. Los conceptos de “creencia” y “deseo” son en este sentido directamente dependientes de las complejidades en la interacción agente / mundo.

Claramente, estas complejidades pueden ser múltiples y extremadamente variadas, dependiendo del nivel de inserción e interacción del agente con su entorno. En este punto reside el carácter gradualista por excelencia de la teoría de los sistemas intencionales, en tanto que hay un alto grado de variabilidad en la “situación ambiental” de un organismo, la cual es afectada por factores tan dispares como los relativos a su historia evolutiva (o su constitución genotípica), a su constitución fenotípica, a su experiencia personal y a la eventualidad del momento de la atribución. Los fenómenos

---

<sup>15</sup> Hay que tener cuidado con no entender aquí que la estrategia intencional sencillamente crea o produce los fenómenos intencionales (lo que llevaría a un interpretacionismo radical), aunque sí puede decirse que los configura, en tanto que el conjunto de patrones que identificamos con el vocabulario mentalista únicamente se nos vuelve inteligible mediante la aplicación de la estrategia (ver pp. 11-12).

intencionales son, de este modo, entendidos como pertenecientes a una clase abierta, como opuestos, digamos, a los términos de clase natural putnamianos<sup>16</sup>. Una implicación gradualista de esto es que el mismo estatus de “sistema intencional” es fundamentalmente no rígido, gradual, por derivación del estatus no rígido de los estados cuya atribución configura un tal sistema.

El carácter gradualista de la teoría de los sistemas intencionales puede entenderse como una consecuencia de la defensa que hace Dennett de lo que voy a llamar la “primacía de la conducta”. ¿Primacía frente a qué? En general, frente a la conciencia y, en particular, frente al rol dado por algunos filósofos de la mente a las representaciones mentales (así como al modo en que el concepto de representación es por ellos entendido)<sup>17</sup>. Aunque no me detendré aquí en el problema de las representaciones mentales, un breve análisis del modo como Dennett enfoca la cuestión pondrá de manifiesto el tipo de gradualismo que quiero destacar.

Al igual que para el caso de los estados intencionales, las representaciones (todo tipo de representación mental) no están constituidas primordialmente por características internas propias de la estructura del cerebro (sean éstas características físicas u operacionales) sino que, otra vez, dependen de la “situación ambiental” del agente y, específicamente, del rol que éstas tienen en la regulación de su conducta. Por este motivo, la diferencia interesante para Dennett no es la que hay entre los sistemas que *realmente* tienen representaciones internas del mundo que los rodea y los que no. El punto es que el funcionamiento de la actitud intencional (y sus posibilidades de éxito explicativo) se mantiene invariante frente a las diversas manifestaciones conductuales, independientemente del estatus de sus representaciones internas.

El gradualismo en la explicación dennettiana de la configuración de los fenómenos intencionales tiene su contraparte semántica en el conjunto de conceptos a los que apelamos para la atribución de contenido mental. Estos conceptos, como máximo, podrán *aludir* de modo más o menos eficaz (y, concordantemente, más o menos preciso) al complejo estado interno de control de la conducta observada, pero su contenido no estará *determinado* por la referencia a algún proceso o etapa de dicho estado.

---

<sup>16</sup> Putnam (1975).

<sup>17</sup> De nuevo, para nuestros propósitos, Fodor. El tema de la primacía de la conducta en el contexto de la disputa entre Dennett y Fodor será un tema central que trataré más adelante (especialmente, en el apartado 4.3).

Ahora bien, no debemos entender que esta explicación es gradualista debido a su carácter perspectivista, que antes destacábamos; es decir, no debe entenderse que las atribuciones intencionales sean esencialmente difusas dada la imposibilidad por parte del intérprete (a causa de su mismo rol como intérprete) de captar de modo exacto el concepto que el agente tiene de su actividad o disposición a actuar. Debemos recordar que la imprecisión ineliminable de las atribuciones intencionales no sólo no comporta problemas para el éxito de la estrategia, sino que este descuido respecto de una determinación conceptual precisa del contenido mental del agente<sup>18</sup> es, en realidad, la característica principal del *mecanismo* por el cual la estrategia funciona, característica a la que por lo tanto deberemos apelar para explicar su éxito; este punto es, en mi opinión, muy relevante y será retomado para algunos desarrollos posteriores.

Me concentraré ahora en el aspecto gradualista metodológico de la teoría dennettiana. Lo primero que quiero destacar es que el carácter perspectivista de la teoría tiene por sí mismo alguna relevancia metodológica. Dennett no asume la categoría de “estado mental” como fenómeno, es decir, como objeto dado a explicar. Dado su perspectivismo, la teoría de los sistemas intencionales no toma como punto de partida un fenómeno supuestamente determinado (por caso, la creencia) para así intentar explicarlo, sino que se centra en el proceso de configuración del fenómeno, esto es, vagamente, nuestra comprensión de sentido común de todo lo relacionado a la mente y la intencionalidad. En este sentido, Dennett parte de una asunción metodológica preliminar de acuerdo con la cual no tenemos de antemano “el” fenómeno de la creencia, tomado así como *factum* a explicar, y por esto los objetivos y alcances de su teoría no se cierran en lo que presumiblemente sería el objeto de estudio sino también en nuestra actitud hacia él. Esto comporta una apertura teórica que es una base para el gradualismo sustantivo de la teoría.

A partir de lo anterior, se entiende que la teoría de los sistemas intencionales no parte tampoco de una delimitación de la gama de agentes hacia la cual habría que dirigirse para investigar los fenómenos mentales: no hay tal delimitación en tanto que no se parte de una preconcepción de dichos fenómenos. En *Kinds of Minds* (1996a, p. 14), Dennett considera que nuestras mentes (las mentes humanas) son necesariamente el baremo a partir del cual debemos empezar la investigación, dada nuestra obvia familiaridad con ellas; pero esto no significa que este tipo de fenómeno (tal como lo

---

<sup>18</sup> Tarea que en *Kinds of Minds* (1996a, p. 56) Dennett denomina “el equívoco objetivo de la precisión proposicional”, asunto al que me dedicaré en detalle en el apartado 4.2.

conocemos de modo familiar) sea excluyente o que la atención exclusiva sobre él podrá ofrecernos su comprensión cabal. Muy por el contrario, Dennett se inscribe dentro del grupo de filósofos que intentan “aclarar primero las necesidades de representación y los recursos de los animales que no utilizan un lenguaje, antes de tratar de elaborar una relación de la creencia humana [...] sobre esa base”<sup>19</sup>. Otra vez, esta actitud no cerrada, en este punto respecto del caso humano en particular, es una precondition para una postura que intenta concebir los fenómenos intencionales como graduales.

Un punto adicional relacionado con el gradualismo de la teoría dennettiana tiene que ver con el mismo estilo filosófico de su autor, un estilo que, como mencionaba al comienzo del capítulo, evade la edificación teórica pretendidamente autónoma y tiende más bien a la formulación de preguntas y a la problematización. Esta manera de asumir la tarea filosófica es en general afín al gradualismo en tanto tiende a justificar el rechazo de preguntas que exigen una respuesta del tipo “todo o nada” o bien, paradigmáticamente, de posiciones que construyen sus objetos teóricos apelando a condiciones necesarias y suficientes.

Para dar un ejemplo relacionado con un tema anterior (aunque no explícito en la obra de nuestro autor), me atrevería a decir que, dada su postura metodológica, no está entre los objetivos teóricos de Dennett la determinación del conjunto (cerrado) de sistemas que entrarían legítimamente en la categoría de “intencionales”: es decir, se subestimarían preguntas del tipo “¿qué animales tienen verdaderamente creencias?”. Y la causa de esto no sería simple falta de interés por el caso, sino que la misma elaboración, por parte de Dennett, de sus herramientas teóricas (en este caso, el concepto de “sistema intencional”) no es compatible con una tal determinación. Todo esto es muy general pero me parece que enmarca a la teoría de Dennett en un cuadro metodológico con un claro carácter gradualista, que se incorpora al carácter gradualista propio de las tesis dennettianas. Paso ahora finalmente a la cuestión del evolucionismo.

---

<sup>19</sup> Dennett (1987a, p. 306).

### ***3. Enfoque Evolucionista acerca del Contenido Mental***

La teoría darwiniana de la evolución por selección natural es la teoría gradualista por excelencia<sup>20</sup>: aporta un modelo de explicación de las características de los organismos vivos (y esto incluye particularmente, para nuestro caso, las que exhiben a través de su conducta) que las entiende como resultados de un proceso constante de variación y selección. La idea gradualista de base aquí es la siguiente: la explicación de las habilidades más sofisticadas y funcionalmente específicas de los organismos, independientemente del grado de esta sofisticación y especificidad, debe en algún punto apelar a otras capacidades de nivel más básico o, si se quiere, a esa misma habilidad entendida en función de características más simples.

Hay, en este sentido, una clara afinidad explicativa general entre los modelos gradualista y evolucionista. Lo que la teoría evolucionista aporta de modo central a una visión gradualista de la mente o del significado es la posibilidad de explicar, desde una perspectiva naturalista, cómo estos fenómenos surgieron, por medio de un proceso generativo gradual compuesto por infinidad de micro-modificaciones, a partir de elementos que carecen de las características definitorias de los primeros. Dado esto, un enfoque evolucionista comportaría la imposibilidad teórica de “trazar líneas nítidas”, esto es, de determinar estrictamente en qué punto del desarrollo de un organismo surge una particular capacidad o característica.

Lo anterior debería dirigir nuestra atención al carácter evolucionista (no explícito) de la apelación, por parte de Dennett, a una estrategia o actitud en la explicación del surgimiento de los fenómenos intencionales (carácter que es coextensivo y coherente con el aspecto gradualista metodológico destacado en el previo apartado). Somos nosotros, de hecho, junto con nuestra adopción de la estrategia intencional, quienes necesitamos “trazar líneas” en tanto que hacer esto nos confiere un poder explicativo y predictivo asombroso, ocultado únicamente por nuestro acostumbramiento a tal ejercicio: nuestra comprensión del mundo está tan impregnada por los modelos intencionales que muy difícilmente nos damos cuenta de su poder y alcance. De este modo, la apelación a nuestra propia perspectiva en tanto que “comprensores del mundo” como punto de partida explicativo puede verse en realidad como una contraparte

---

<sup>20</sup> Esta simple idea en realidad tiene sus oponentes, uno de los cuales es Stephen Jay Gould. Como no es mi objetivo aquí dar una defensa exhaustiva de la relación entre darwinismo y gradualismo, remito el lector a la discusión dennettiana de las objeciones de Gould, en el décimo capítulo de *Darwin's Dangerous Idea* (1995).

obligada de una teoría del contenido mental que atribuya un rol importante a la teoría evolucionista.

Otro aspecto gradualista de la teoría dennettiana que podemos reformular aquí en clave evolucionista es la atención prestada a los animales no lingüísticos. Hay por lo menos una presunción de que este foco de atención nos proporcionará claves para la comprensión de los problemas particulares propios del caso humano. Esto es justificable en tanto se asuma el principio evolucionista de la necesaria continuidad en la conformación de las habilidades de los organismos, así como de las características a las que deberemos apelar para su explicación. Lo que este principio implica es que habrá muchos aspectos en que nuestras mentes se asemejarán a las de otros animales y muchos otros en los que diferirán, aunque también en este caso será posible y explicativamente necesario rastrear las fuentes evolutivas de estas diferencias. Este es un factor adicional por medio del cual puede interpretarse que Dennett reconstruye los problemas filosóficos de la intencionalidad desde la perspectiva metodológica gradualista / evolucionista<sup>21</sup>.

Los puntos anteriores pueden verse más bien como consecuencia de la influencia metodológica que comporta la incorporación de la teoría evolucionista en el planteo filosófico dennettiano. Por otro lado, hay también puntos cruciales del planteo en el que se apela directamente a la explicación evolucionista. En *Kinds of Minds*, Dennett presenta explícitamente el enfoque teórico de su postura como evolutivo, mostrando la lenta genealogía del concepto de mente, desde sus orígenes más sencillos (en el que se nos haría muy difícil hablar estrictamente de “mente”) hasta llegar al fenómeno infinitamente más complejo con el que asociamos normalmente el término. Esta relación cercana con el evolucionismo puede rastrearse en parte en lo que William Lycan (1981) ha llamado el “Homuncularismo” de la teoría de los sistemas intencionales.

El homuncularismo es la idea de que la única explicación disponible para el conjunto de fenómenos relacionados en general con el comportamiento inteligente es su progresiva descomposición en elementos cada vez menores y, concordantemente, cada vez menos inteligentes; estos elementos pueden entonces ser vistos, adoptando la estrategia intencional, como micro-agentes (homúnculos) que responden a sus

---

<sup>21</sup> De aquí en adelante daré por supuesto el vínculo estrecho que puede mantenerse entre la incorporación de consideraciones evolucionistas y la defensa de una explicación de tipo gradualista: esto es, toda vez que se traten argumentos evolucionistas se estará implicando su gradualismo.

necesidades individuales actuando de uno u otro modo. Dado el gradualismo evolucionista, un modelo apropiado para entender la conformación de una mente, tal como utilizamos el término desde el sentido común (esto es, aplicado a organismos de comportamiento flexible y complejo), es justamente apelando a una continua complejización de las actividades de elementos máximamente “estúpidos” que exhiben un comportamiento simple y mecánico. Esta es una manera metafórica de expresar la idea de que en definitiva “¡estamos compuestos de robots!” y que el origen de toda forma compleja de intencionalidad puede finalmente rastrearse en las actividades mecánicas proto-intencionales de millones y millones de elementos atómicos organizados en sistemas.

Un modelo diferente propuesto por Dennett, y que también apela a la idea de una evolución de la intencionalidad para explicar las manifestaciones más complejas del fenómeno, es lo que en el artículo “The Role of Language in Intelligence” (1994) el filósofo llama la “Torre de la Generación y la Prueba”<sup>22</sup>. Este es un relato evolutivo que aporta una estructura simplificada de las posibles encarnaciones del proceso que va desde formas simples de intencionalidad hasta el caso ultra-especializado de la raza humana; es decir, un relato que identificaría los modos generales de organización de los micro-agentes homunculares que recién mencionábamos.

Es importante tener en cuenta que, al igual que para el caso del homuncularismo, esta propuesta no tiene pretensión alguna de describir procesos fisiológicos particulares, sino que su utilidad es aportar un marco para dar sentido a la idea de una evolución de la intencionalidad. Este es un punto que se relaciona con el tema, en mi opinión fundamental, de entender precisamente cuál es el rol de (algunas de) las ideas aportadas por Dennett y cuál sería su relación con el estudio psicológico. Una gran cantidad de críticas y malas interpretaciones podrían evitarse si se lograra encuadrar la teoría dennettiana dentro de los propósitos para los que fue propuesta. Esta es una tarea que me parece importantísima, y a la que trataré de contribuir en la conclusión de este trabajo.

Dennett subdivide la torre de la generación y la prueba en cuatro etapas, cada una representada por un tipo de organismos: las criaturas darwinianas, las criaturas skinnerianas, las criaturas popperianas y las criaturas gregorianas. Los diferentes pisos

---

<sup>22</sup> El modelo de la torre de la generación y la prueba es también presentado en *Darwin's Dangerous Idea* (1995, pp. 373 y siguientes) y en *Kinds of Minds* (1996a, pp. 103 y siguientes), e incluso pueden encontrarse múltiples referencias a él en varios otros puntos de la obra dennettiana.

de la torre se construyen uno sobre la base del anterior e implican un avance en el poder cognitivo del organismo, aunque esto no significa que los modelos previos sean reemplazados por los siguientes, sino que la relación es complementaria.

Las criaturas darwinianas exhiben fundamentalmente el rasgo básico del proceso evolutivo, esto es, la selección de un fenotipo estructural en base a su capacidad para la supervivencia en su medio ambiental. En este caso el fenotipo seleccionado es fijo y se conserva por transmisión genética. El paso hacia las criaturas skinnerianas está marcado por su adquisición de la propiedad de la “plasticidad fenotípica” (*phenotypic plasticity*). Esto significa que el diseño estructural de estos organismos no es fijo, sino que puede modificarse en base al impacto ambiental sobre ellos. Estas criaturas probarían diferentes acciones y algunas de ellas tendrían la suerte suficiente para seleccionar las acciones favorables para su subsistencia. Esta “selección” únicamente es debida a la presencia de lo que Dennett llama “reforzadores” (*reinforcers*), mecanismos internos con la capacidad de favorecer mecánicamente un determinado tipo de acciones, es decir, actuar sobre otros mecanismos de control de la conducta aumentando su actividad. Los organismos nacidos con refuerzos adecuados para su supervivencia serán seleccionados.

Un avance considerable respecto de esta selección azarosa es la posibilidad de *preseleccionar* una de varias opciones disponibles para la acción, antes de ponerlas en práctica corriendo así riesgo. Las criaturas popperianas están de este modo dotadas de la posibilidad de albergar internamente información sobre el entorno externo y los diferentes resultados de una interacción con él. El paso hacia las criaturas gregorianas está finalmente dado por la posibilidad de utilizar herramientas (esto es, porciones diseñadas del entorno externo) para la manipulación de la información en el entorno interno que permite la preselección. El aporte dado por el uso de herramientas es la posibilidad de una manipulación más efectiva de la información sensorial disponible, además del hecho considerable de obtener beneficio de los avances hechos por otros organismos en la elaboración de herramientas siempre más eficientes<sup>23</sup>.

Hay una derivación meta-teórica que puede extraerse del carácter evolucionista de los modelos homuncularista y de la torre de la generación y la prueba. Me interesa especialmente mostrar la analogía explicativa entre el ya destacado continuismo de la perspectiva homuncularista y el paso crucial que se da, en el segundo modelo, entre las criaturas skinnerianas y las popperianas. El punto es el siguiente: la postulación de un

---

<sup>23</sup> Los problemas particulares relativos a las criaturas gregorianas y a la utilización de herramientas serán el tema central del apartado 3.2. No me detendré aquí en ellos.

ambiente interno para el control conductual en las criaturas popperianas no implica la necesidad de invocar mecanismos físicos esencialmente diferentes de las rígidas relaciones de refuerzo de la conducta, claras en los casos de las criaturas darwinianas y skinnerianas. El principio de funcionamiento de esas relaciones es la llamada “Ley del Efecto”, principio fundamental típico de las explicaciones del conductismo ortodoxo (skinneriano), según el cual toda acción seguida de recompensa será repetida.

Dennett (1975) mantiene que la ley del efecto no puede abandonarse para la elaboración de cualquier teoría adecuada del comportamiento, y en este punto reside el valor meta-teórico compartido por el homuncularismo y la torre de la generación y la prueba, dado el carácter evolucionista de ambos modelos. En efecto, Dennett concibe la ley del efecto no tanto como análoga al principio de la selección natural sino más bien como una continuación suya (lo que podría verse reflejado en el paso de las criaturas darwinianas a las skinnerianas):

La Ley del Efecto presume que hay una “población” de pares de estímulo y respuesta, unidos más o menos fortuitamente o en todo caso arbitrariamente, y desde este grande y variado conjunto, los reforzadores seleccionan los pares bien diseñados, adaptados, fortuitamente apropiados, de modo *enteramente mecánico*<sup>24</sup>.

La idea continuista del homuncularismo dennettiano se refleja en la extensión entre los principios evolucionista (selección natural) y conductista (ley del efecto) para todos los estadios de la torre de la generación y la prueba.

Los modelos aquí destacados (el homuncularismo y la torre de la generación y la prueba) se hacen eco de la teoría evolucionista para presentar una visión gradualista de la conformación de las manifestaciones más sofisticadas de lo que llamamos mente o de su intencionalidad. Hay, por otra parte, un punto fundamental en la teoría de los sistemas intencionales que establece un vínculo *conceptual* entre la teoría darwinista y la noción de “sistema intencional”. Este punto es a la vez lo que explica la apelación al trascurso evolutivo de un organismo en la determinación de lo que llamábamos (p. 9) su “situación ambiental”. El vínculo conceptual en cuestión es el que se da entre el proceso evolutivo y nuestras asunciones de racionalidad, propias de la estrategia intencional: esto es, dado el origen evolutivo de un organismo, podemos suponer con seguridad que la mayoría de sus creencias serán verdaderas y que, frente a una situación determinada, su actuar en base a tales creencias será consecuente con sus objetivos.

---

<sup>24</sup> Dennett (1978a, p. 73). La traducción y el subrayado son míos.

Ya en su primera obra, *Content and Consciousness*, Dennett reconoce esta relación implícita entre racionalidad y evolución<sup>25</sup>:

El vínculo implícito entre cada trozo de interpretación intencional y su base extensional es una hipótesis o serie de hipótesis que describen la fuente evolutiva del arreglo fortuitamente beneficioso *en virtud del cual* la operación del sistema tiene, en esta instancia, sentido<sup>26</sup>.

La atribución de racionalidad sigue siendo una asunción explicativamente conveniente, pero su adecuación (y, por lo mismo, su éxito) *únicamente* puede retrotraerse hasta el desarrollo evolutivo del agente.

Es importante entender que esta relación entre la asunción de racionalidad y el evolucionismo constituye un vínculo profundo; es decir, no apelamos de hecho (mucho menos, explícitamente) a la teoría evolucionista cada vez que atribuimos contenido desde la actitud intencional<sup>27, 28</sup>. Con palabras de Dennett, la idealización implícita en el mecanismo de la actitud intencional “nos da la noción de un operador o agente epistemológico conativo ideal [esto es, idealmente racional], reducido a un conjunto de necesidades para la supervivencia y la procreación y al entorno en el que evolucionaron sus antepasados y al cual está adaptado”<sup>29</sup>.

La raíz evolucionista del concepto de racionalidad se encuentra en el hecho de que el proceso evolutivo es, inexorablemente, la fuente última de todo objetivo, para cualquier sistema intencional. De nuevo, esta es una tesis filosófica, absolutamente general, que no pretende explicar los detalles de los actos cotidianos de atribución, sino inquirir en las raíces conceptuales, en este caso, de la idea de un “ser racional”. En cuanto que filosóficos, estos problemas están íntimamente relacionados con el debate (estrictamente filosófico) entre los defensores de una intencionalidad intrínseca u original y aquellos quienes niegan la pertinencia de un tal concepto. En lo que sigue quisiera presentar brevemente la posición dennettiana al respecto.

Dennett trata puntualmente el mencionado debate en “Evolution, Error and Intentionality” (1987b), donde, a través de diversos experimentos mentales, trata de

---

<sup>25</sup> A pesar de esto, Dennett (1987a, p. 94) explícitamente niega la posibilidad de dar una definición de la racionalidad en términos evolucionistas y, más aún, da razones para obviar la pregunta por la naturaleza de lo racional.

<sup>26</sup> Dennett (1969, p. 114). El subrayado es mío.

<sup>27</sup> Es por esto también que la actitud intencional puede extenderse hacia todos los sistemas *no* biológicos para los cuales aporte poder predictivo y explicativo, aunque el proceso de desarrollo al que haríamos alusión no será evolutivo.

<sup>28</sup> Esta es una crítica que de hecho Jonathan Bennett (1993, p. 383) formula contra Dennett.

<sup>29</sup> Dennett (1987a, p. 55).

mostrar el equívoco que hay en mantener una distinción entre dos tipos esencialmente diferentes de intencionalidad, la original, paradigmáticamente exhibida por nosotros los seres humanos, y la derivada, propia de nuestros artefactos en general cuya intencionalidad justamente es tal únicamente en tanto depende de una derivación de nuestro tipo particular de intencionalidad.

El argumento básico de Dennett para mostrar esto apela al proceso evolutivo entendido, como veíamos, como la fuente irreducible de todo propósito y, menos metafóricamente<sup>30</sup>, de toda función en y para el organismo. Dado esto, el argumento se completa con la defensa de un vínculo estrecho entre función y significado, en tanto que todo significado es tal únicamente si es puesto en relación con el esquema mayor de cosas dentro del que opera<sup>31</sup>. Ahora, la función no es una propiedad irreducible o fija, sino que su determinación y especificación puede rastrearse a lo largo del lento proceso evolutivo que, modificación tras modificación en las características relevantes del organismo, fue constituyéndola: este proceso es, finalmente, el esquema de cosas último hasta el cual podemos derivar toda función y, por lo mismo, todo significado, toda intencionalidad.

Resulta claramente contraintuitivo decir que nuestra intencionalidad (esto es, nuestra capacidad para el significado) es derivada de la intencionalidad del proceso de evolución por selección natural, un proceso una de cuyas características claves es la arbitrariedad de las modificaciones sobre las que opera. Dennett reconoce este punto y aporta una explicación para él, estrechamente relacionada con los motivos por los que entiende que el adaptacionismo<sup>32</sup> en biología es una estrategia de estudio absolutamente continua con la vertiente intencionalista<sup>33</sup> en psicología. Dennett entiende que lo que hay de contraintuitivo en hablar de la intencionalidad del proceso evolutivo (como algo tan real como nuestra propia intencionalidad) se debe a lo que denomina “chovinismo de escala temporal”, que ilustra con un experimento mental:

[S]i nuestro planeta fuera visitado por marcianos que pensarán los mismos pensamientos que nosotros sólo que miles o millones de veces más rápidamente, les pareceríamos tan idiotas como los árboles y

---

<sup>30</sup> O, como veremos, en definitiva, de modo igualmente metafórico.

<sup>31</sup> Este es el sentido en que la teoría dennettiana del contenido es funcionalista; profundizaré sobre esto en la segunda parte del apartado 4.3.

<sup>32</sup> La idea de que es metodológicamente justificado asumir la optimidad del diseño de los organismos y sus estructuras funcionales.

<sup>33</sup> Esto es, simplemente, la consideración de que es justificado para explicar los procesos psicológicos utilizar el vocabulario de intenciones y propósitos de la psicología de sentido común.

tendrían la propensión a mofarse de la hipótesis de que nosotros tenemos mente<sup>34</sup>.

La intencionalidad de la selección natural es perfectamente real, pero su operación no es fácilmente discernible debido a la escala temporal en la que se desenvuelve. En este sentido acotado, estamos exactamente tan justificados para hablar no metafóricamente de nuestra intencionalidad como lo estamos para hablar del mismo modo de la intencionalidad de la “Madre Naturaleza”. La estrategia a través de la cual los patrones intencionales son individuados (esto es, la actitud intencional) funciona a través de los mismos principios y por el mismo mecanismo cuando versa sobre los procesos psicológicos así como cuando lo hace sobre el proceso evolutivo a gran escala.

Otro aspecto de la relación que Dennett mantiene entre la actitud intencional y el proceso evolutivo es la idea (aunque no muy desarrollada en su obra) de la estrategia intencional como arraigada evolutivamente. Esto debe entenderse en un sentido muy concreto. La estrategia intencional, como característica fundamental de un conjunto importante de nuestros emprendimientos cognitivos, nos otorga una clara ventaja a la hora de comprender el mundo para su manipulación. En este sentido, puede entenderse la adopción de la estrategia como un elemento constitutivo de nuestras capacidades cognitivas y, a la vez, un elemento adaptativo (o reforzado) dado el efecto beneficioso que comporta en los resultados de esas capacidades.

Esto, a la vez, da pie para interpretar la actitud intencional como la manifestación de un conjunto de características básicas, propias del organismo capaz de adoptarla. Estas ideas pueden verse como un argumento en contra de quienes interpretan el discurso mentalista de la psicología popular esencialmente como una teoría, cuya lógica interna se debería a nuestro esfuerzo conciente de comprender el mundo. Por lo menos, dada la aducida raíz evolutiva de la actitud intencional, habría razones para buscar elementos no aprendidos en la psicología de sentido común, y para entenderla no tanto como una teoría sino más bien, y de acuerdo con Dennett, como una habilidad arraigada en el comportamiento de algunos organismos<sup>35</sup>.

Es importante destacar, finalmente, la medida en que Dennett justifica hablar de la evolución de la intencionalidad o de la mente. Este es, para él, un discurso explicativamente no negociable en su teoría del contenido. Toda atribución de contenido

---

<sup>34</sup> Dennett (1996a, p. 78).

<sup>35</sup> Este punto será extensamente desarrollado en el apartado 4.1, ya que será central para mi discusión comparativa de las teorías dennettiana y fodoriana.

mental, dadas sus apelaciones de optimidad o racionalidad, debe necesariamente encuadrarse en un nivel general de interacción, desarrollo y evolución entre el sistema interpretado y el entorno en que éste actúa. Por este punto en particular, además de todo lo dicho anteriormente en este apartado, sea por las afinidades metodológicas como por las apelaciones dennettianas explícitas a la teoría evolucionista, considero justificado encuadrar a la teoría de los sistemas intencionales en lo que puede llamarse el *enfoque evolucionista acerca del contenido mental*. Lo llamo “enfoque” porque entiendo que dicha teoría asume de modo central una perspectiva evolucionista, que influye en definitiva en todos los aspectos y los problemas relativos a la comprensión de la mente. En el siguiente capítulo me dedicaré a presentar lo que, según defenderé, es un representante ilustre del enfoque exactamente opuesto, esto es, la Teoría Representacional de la Mente, propuesta por Jerry Fodor.

---

---

## La Teoría Representacional de la Mente

### *1. Presentación General de la Teoría*

En este apartado presentaré la llamada “Teoría Representacional de la Mente” de Fodor, limitándome a sus aspectos más pertinentes para la confrontación con la posición dennettiana. Quiero aclarar nuevamente que mi interés en este trabajo está centrado principalmente en algunos aspectos de la teoría de los sistemas intencionales y que, en este sentido, el análisis de la postura de Fodor será instrumental y su alcance estará limitado por su utilidad para el contraste con dicha teoría. De este modo, por un lado, algunas temáticas muy elaboradas por Fodor, como la de la arquitectura modular de la mente / cerebro y la de su atomismo conceptual y su nativismo extremo, aunque en alguna medida relevantes para su teoría del contenido, serán descartadas en tanto que Dennett no se ocupa de ellas y, por otro lado, algunas discusiones técnicas sobre puntos específicos dentro de dicha teoría, como el “problema de la disyunción” (pp. 39-40), no serán desarrolladas a fondo en tanto que desde la perspectiva dennettiana se critica su misma pertinencia. Al igual que en el capítulo anterior, mi interés versará en el relevamiento del marco metodológico general en el que, en este caso, Fodor se circunscribe para elaborar su postura. El estudio de la teoría fodoriana en estos términos nos permitirá justificar finalmente su inclusión en lo que llamaré “enfoque anti-evolucionista de lo mental”.

A diferencia de Dennett, Fodor es un filósofo estructurado, rígido y fiel al método argumentativo tradicional de análisis de premisas y derivación de conclusiones. La exposición de las tesis defendidas es casi siempre ordenada, explícita y exhaustiva. Dado esto, encuentro conveniente para la presentación de la teoría fodoriana seguir una de sus varias formulaciones canónicas, específicamente la última de ellas, que da comienzo a *Concepts* (1998). Usaré esta formulación como eje para la exposición, pero en ningún modo me limitaré a ella: mi análisis abarcará buena parte de la producción filosófica fodoriana pertinente, y en especial *The Language of Thought* (1975),

*Representations* (1981), *Psychosemantics* (1987) y *The Elm and the Expert* (1994), además de otros artículos relevantes.

Cabe aclarar que un problema importante para el presente capítulo ha sido el de integrar las tesis nucleares de la teoría, teniendo en cuenta algunas modificaciones y revisiones sustanciales que ha sufrido sobre todo en los últimos años (notablemente, el abandono de la noción de contenido estrecho en *The Elm and the Expert* y el redimensionamiento general de la teoría en *The Mind Doesn't Work That Way* del año 2000). De todos modos, por una parte, la estructura y los puntos básicos de la teoría se siguen manteniendo y, por otra parte, mi interés apunta a sus aspectos centrales y al modo cómo éstos pueden delinear una perspectiva general para el abordaje de los problemas referidos a la intencionalidad y el contenido mental: en este sentido, tomaré la teoría como, con palabras de Fodor, “una vaga confederación de tesis”<sup>36</sup>.

La primera tesis de la teoría representacional de la mente, según la formulación canónica de *Concepts*, es la más importante para mis propósitos ya que refiere en parte a la interpretación de la psicología de sentido común y el papel que juega para el desarrollo de una psicología científica. La tesis es la siguiente:

*La explicación psicológica es típicamente nómica y es completamente intencional. Las leyes que invoca la explicación psicológica expresan, de manera típica, relaciones causales entre estados mentales que se especifican bajo una descripción intencional; es decir, entre estados mentales que se escogen por referencia a sus contenidos*<sup>37</sup>.

Esta tesis (a la que me referiré como T1) es más compleja de lo que puede parecer ya que acarrea muchos supuestos centrales de la teoría fodoriana; por esto, me concentraré especialmente en ella.

En una primera aproximación, T1 versa esencialmente sobre la “explicación psicológica”, cuyas dos características principales son la de ser nómica, esto es, sus generalizaciones son susceptibles de figurar en leyes científicas, y la de ser intencional, esto es, tales generalizaciones son especificadas en los términos intencionales mentalistas típicos de nuestro discurso psicológico de sentido común. Estas dos ideas en conjunto se traducen en la tesis fuerte de la existencia de *leyes intencionales* y, por lo tanto, de la posibilidad de una ciencia psicológica especificada en vocabulario intencional. Ahora, aunque el énfasis metodológico de Fodor está puesto, a lo largo de toda su trayectoria, sobre el carácter empírico / científico de la investigación sobre la

---

<sup>36</sup> Fodor (1998, p. 24).

<sup>37</sup> *Idem.*

mente, T1, que puede entenderse como el fundamento de su teoría de la intencionalidad, establece una conexión fuerte entre dicha investigación y la psicología popular, conexión que delinea para esta última una serie de características importantes. En lo siguiente me concentraré en este punto.

Al igual que para el caso de Dennett, Fodor toma el discurso mentalista de la psicología popular como punto de partida para su teoría de la intencionalidad. A pesar de que, como veremos en el capítulo 4, tanto las motivaciones como las consecuencias teóricas de la interpretación de dicho discurso sean diametralmente opuestas para ambos autores, puede delinearse un conjunto de sus características, relacionadas con el carácter y la importancia de la psicología popular, que ambos filósofos defienden: es decir, este conjunto de características puede mantenerse aunque sirvan de trasfondo a teorías altamente incompatibles.

Un primer punto a destacar es el poder predictivo de la psicología popular, un poder que damos a menudo por sentado, considerando la frecuencia con que usamos la explicación intencional: ésta funciona independientemente del conocimiento de las personas (u organismos) cuyo comportamiento queremos predecir y del conocimiento de las variables que puedan interactuar en su situación. El éxito de las predicciones intencionales es extremadamente alto y es frustrado normalmente sólo en los casos en que en la situación interviene alguna variable inesperada que la predicción no pudo haber contemplado. Otra característica compartida es la indispensabilidad de la psicología. Se entiende que la psicología popular se sitúa en un nivel explicativo cuyo rol no puede concebiblemente ser ocupado por un discurso diferente. La idea, en contra de eliminativistas como Stephen Stich (1983) y Paul Churchland (1981), es que no podemos abandonar la psicología popular porque sencillamente no sabríamos cómo hacerlo: no sabríamos cómo explicarnos mutuamente la conducta apelando a términos no intencionales, no mentalistas.

El nivel explicativo intencional es entonces no negociable. Esto tiene una derivación a tener en cuenta, que es resaltada también por Dennett. Las explicaciones y predicciones intencionales constituyen en conjunto un nivel teórico cerrado, distintivo, con características particulares y con un vocabulario propio. Desde un enfoque programático en vista a un estudio científico del comportamiento inteligente, en sus varias manifestaciones y desde las diferentes perspectivas científicas pertinentes, tanto Fodor como Dennett ven la necesidad de preservar el tipo de explicación que utiliza el

vocabulario intencional<sup>38</sup>. Hay por parte de ambos filósofos un rechazo explícito de una posible reducción explicativa del nivel intencional al nivel más básico de las ciencias físicas: todo lo que pueda decirse meramente en términos de los movimientos de los organismos no agotaría el poder explicativo adquirido en términos de creencias y deseos.

Volviendo ahora nuestra atención a otros aspectos importantes de la concepción fodoriana de la psicología popular, pero divergentes respecto de la concepción dennettiana, el aspecto central a tener en cuenta es la analogía que Fodor mantiene entre el *tipo* de explicación intencional de sentido común y la explicación científica: otros aspectos de la concepción fodoriana derivan más o menos directamente de éste. La idea básica consiste en entender la psicología de sentido común como una suerte de teoría en lo que refiere a su estructura y al tipo de generalizaciones que con ella realizamos, aunque éstas sean menos estrictas y menos precisas que las realizadas en las teorías científicas:

Hay [...] una franca similitud entre la manera en la que las generalizaciones implícitas funcionan en la psicología de sentido común y la manera en la que las generalizaciones explícitas funcionan en las ciencias especiales<sup>39</sup>.

Más aún, Fodor sostiene que lo que llama la “estructura deductiva” propia de las ciencias está frecuentemente presente en las generalizaciones explícitas de la psicología popular. Y esto por dos razones: por un lado, porque las generalizaciones intencionales apelan a no observables y, por otro lado, porque sus predicciones se obtienen cuando las generalizaciones se repiten e interactúan (1987, p. 24).

Dado el carácter teórico del discurso mentalista de sentido común, Fodor puede pasar a sostener la tesis meta-teórica fuerte de que el único formato adecuado para una psicología científica que dé cuenta de las propiedades intencionales de la mente es el de una teoría que postule estados mentales del tipo que comúnmente nos atribuimos para explicar y predecir nuestro comportamiento, esto es, estados como las llamadas “actitudes proposicionales”<sup>40</sup>. Baste con decir por ahora que las actitudes proposicionales pueden ser de diferentes tipos (creencia, deseo, intención, esperanza, temor, etcétera) y se identifican por su contenido. De acuerdo con esto, los estados

---

<sup>38</sup> Aunque, como se discutirá en el capítulo 4.4, las motivaciones son diferentes en cada caso y, de acuerdo con esto, el tipo de explicación considerada posible y pertinente diferirá fuertemente.

<sup>39</sup> Fodor (1987, p. 22).

<sup>40</sup> Este término fue propuesto primeramente por Bertrand Russell (1940). Discutiré las implicaciones de adoptar este modelo para los estados mentales en el apartado 4.2.

postulados por las teorías científicas deberán corresponderse aproximadamente con la tipología de actitudes que el discurso cotidiano recoge y a la vez serán individuados por el contenido proposicional que les es propio. Las leyes de una futura psicología científica completa son de este modo entendidas como una continuación y elaboración de las generalizaciones de sentido común en las que nos apoyamos cotidianamente: en otras palabras, la tarea de la psicología cognitiva es la de sistematizar y hacer explícitas y rigurosas el mismo tipo de generalizaciones que son propias de la psicología de sentido común.

Un paso ulterior que da Fodor es establecer una relación de justificación epistémica entre las leyes psicológicas así concebidas y el discurso de la psicología popular. La tendencia a interpretar este discurso como un discurso esencialmente teórico genera el marco para plantear la cuestión de su adecuación *qua* teoría. La idea entonces es que la elaboración de una teoría psicológica verdadera que utilice el mismo tipo de generalizaciones (la misma “estructura deductiva”) que las de la psicología popular y que construya sus entidades teóricas apoyándose en las ya utilizadas por la psicología popular, funcionará retrospectivamente como justificatoria respecto de ésta: es decir, como una teoría que corrobore la verdad general de las generalizaciones toscas e implícitas de sentido común. Con palabras de Fodor:

*Continuar con las actitudes – reivindicar la psicología de sentido común – significa mostrar cómo se podría tener (o, como mínimo, mostrar que uno podría tener) una ciencia respetable cuya ontología reconociera explícitamente estados que exhiben la clase de propiedades que el sentido común atribuye a las actitudes<sup>41</sup>.*

Además de su carácter nomológico e intencional, las otras dos características señaladas por T1 refieren a que las leyes intencionales, por un lado, especifican relaciones causales entre estados mentales y, por otro lado, como mencioné, que estos últimos son identificados por su contenido. Estas características son extremadamente importantes para el planteo fodoriano en general y especialmente respecto del modo cómo en él el discurso psicológico popular se interrelaciona con el científico, tanto que en algunas ocasiones son presentadas incluso como *condiciones* básicas para que una psicología de sentido común pueda ser considerada como tal (por ejemplo, Fodor, 1987, p. 29). La primera característica es expresada por la condición de que los estados postulados por la psicología de sentido común tengan poderes causales y la segunda por

---

<sup>41</sup> Fodor (1987, p. 28). El subrayado es mío.

la condición de que los mismos sean semánticamente evaluables. Me volveré hacia cada una de ellas respectivamente, antes de pasar a la segunda tesis de la teoría representacional de la mente, T2.

De acuerdo con la primera característica, cuando cualquiera de nosotros dice que hará algo *porque* tiene determinadas creencias y deseos, puede interpretarse correctamente la justificación (el “porqué”) de la acción como causal en algún sentido. Se entiende luego que la explicación psicológica de sentido común está comprometida con (es decir, asume) la *causación mental*. Esta puede ser de tres tipos diferentes: la causación del comportamiento por los estados mentales, la causación de los estados mentales por la estimulación ambiental y la causación entre los mismos estados mentales. Este último tipo de causación es la que posibilita el fenómeno del razonamiento, entendido como la concatenación lógicamente estructurada de pensamientos, entendidos a la vez como especies de sucesos mentales complejos. El que los estados mentales tengan poderes causales es un requisito necesario para que la psicología que apele a ellos pueda ser empírica, dada la relación estrecha que según Fodor una psicología tal mantendría con el discurso intencional de sentido común.

El que los estados mentales sean evaluables semánticamente equivale a sostener que su individuación en tanto estado es determinada apelando al contenido que les es propio. Otra manera de expresar el requisito de evaluabilidad semántica es decir que las creencias y deseos a los que apelamos cotidianamente tienen condiciones de satisfacción. Por ejemplo, en el caso de una creencia, sus condiciones de satisfacción serían el estado de cosas en el mundo en virtud del cual sería verdadera o falsa y, en el caso de un deseo, sus condiciones de satisfacción serían el estado de cosas en el mundo en virtud del cual se vería satisfecho o frustrado. El contenido de un estado mental es, en este sentido, lo que lo hace ser acerca de algún aspecto del mundo.

Ahora bien, los estados mentales poseen su contenido de un modo *constitutivo*: parte de lo que es ser un estado mental determinado es poseer un contenido intencional determinado<sup>42</sup>. Más aún, la noción del contenido de un estado psicológico adquiere un rol central en la teoría fodoriana, no sólo para una definición de lo que se entiende por “estado mental” en tanto que noción clave para una psicología cognitiva sino también como eje del tipo de explicación al que en ella se recurre: esto es, una explicación que subsume los estados mentales por referencia a sus contenidos intencionales.

---

<sup>42</sup> Esta idea será desarrollada ulteriormente en la discusión de T2.

En la psicología popular, ambas características están estrechamente vinculadas. Fodor considera que uno de los hechos fundamentales de la “mente cognitiva” es que las relaciones causales entre los estados mentales respetan característicamente las relaciones semánticas entre los contenidos de los mismos; es decir, que existe un paralelismo entre las relaciones causales y semánticas en las que el discurso mentalista de sentido común se apoya. Aceptadas ambas características y aceptada su interrelación, la teoría representacional de la mente es propuesta como una explicación de cómo podría haber estados con las propiedades causales y semánticas que el sentido común supone en la explicación intencional. El paralelismo entre relaciones causales y semánticas de los estados mentales es, así, el punto nuclear del *explicandum* de la teoría.

La teoría representacional de la mente es un exponente eminente y explícito del realismo acerca de los estados intencionales (también llamado por Fodor “realismo acerca de las actitudes proposicionales”), y justamente a partir de las características destacadas y de su interrelación. Fodor (1985) propone considerar que un realista acerca de las actitudes proposicionales sostiene, por una parte, que hay estados mentales cuya ocurrencia e interacciones causan el comportamiento y, por otra parte, que estos estados son semánticamente evaluables. De acuerdo con esto, el realismo acerca de los estados intencionales sostiene la existencia de estados mentales entendidos como eventos concretos con las características asumidas por nuestro discurso mentalista cotidiano: estados que poseen esencial o constitutivamente propiedades intencionales, propiedades que a la vez están causalmente implicadas en la producción del comportamiento. El realismo fodoriano es aún más fuerte si le agregamos la tesis antes considerada de que una psicología científica necesariamente operará sobre los estados mentales así concebidos, respetando las asunciones implícitas de las generalizaciones de sentido común.

Paso ahora a T2, que sostiene lo siguiente: “[l]as “representaciones mentales” son las portadoras primitivas del contenido intencional”<sup>43</sup>. T2, por una parte, introduce el concepto clave de “representación mental” y, por otra parte, puede considerarse una formulación de la tesis de la intencionalidad intrínseca (tesis que ya presenté en pp. 23-24), en la versión representacionista fodoriana. A medida que el debate respecto de la posibilidad de una intencionalidad original o intrínseca fue desarrollándose (sobre todo a partir del artículo de John Searle de 1980, “Minds, Brains and Programs”, donde se

---

<sup>43</sup> Fodor (1998, p. 24).

presenta el problema), Fodor fue siendo más explícito al respecto. A pesar de no haber sido defendida como tesis en otras presentaciones de la teoría representacional de la mente, T2 siempre fue uno de sus supuestos centrales. Ahora, antes de considerar la tesis fodoriana de la intencionalidad intrínseca, deberé presentar el concepto al que ésta refiere, esto es, el de representación mental, verdadero núcleo de la teoría de Fodor.

La teoría se centra en una noción fuerte de representación mental que cumple en ella diversos roles. Las representaciones mentales son entendidas básicamente como entidades instanciadas físicamente en el cerebro y cuya instanciación es una condición necesaria para que un organismo se encuentre efectivamente en un estado mental determinado: es decir, concretamente, “no hay un episodio de creer-que-P [desear-que-P, temer-que-P, etcétera] sin el episodio correspondiente de una instancia-de-una-representación-mental”<sup>44</sup>.

Los estados intencionales son de este modo concebidos como estados relacionales entre un organismo y representaciones mentales. Las distintas relaciones que el organismo mantiene con las representaciones constituyen los diferentes tipos de actitudes proposicionales (creencia, deseo y demás). Fodor es funcionalista acerca de la individuación de los estados en lo que refiere a su clase o tipo; es decir, lo que hace que un estado mental particular sea una creencia, y no un deseo u otra clase de estado, es el rol funcional que ocupa (como *output* en relación con la estimulación sensorial, como *input* para otros procesos cognitivos y en su interacción con otros estados mentales).

Dije más arriba que la noción fodoriana de representación mental es fuerte no sólo porque se requiere que las representaciones sean instanciadas en estructuras físicas (neuronales) del cerebro, sino también porque se requiere que sean simbólicas y, más específicamente, que posean ellas mismas una estructura *oracional*: esto es, las representaciones mentales son símbolos y, como tales, poseen propiedades sintácticas (o formales) y semánticas. Esta es la famosísima (y discutidísima) tesis del lenguaje del pensamiento (Fodor, 1975), lenguaje también denominado, en algunas ocasiones, “mentalés”.

De acuerdo con esta tesis, las representaciones mentales, al igual que los lenguajes naturales, poseen una sintaxis combinatoria y una semántica composicional. Dado el primer punto, las representaciones pueden combinarse creando representaciones más complejas que tienen a las primeras como sus constituyentes; dado el segundo

---

<sup>44</sup> Fodor (1998, p. 26).

punto, el contenido de las representaciones más complejas está en función del contenido de las representaciones constituyentes y el modo en que se relacionan entre ellas. A la vez, la sintaxis es considerada por Fodor (1987, p. 40) una propiedad física de orden superior: esto es, una propiedad abstracta de su forma, con lo que el requisito de instanciación física de las representaciones puede configurarse de acuerdo con la tesis del lenguaje del pensamiento. Dicho de otro modo, las representaciones mentales sólo pueden afectar a un sistema físico, y por ende sólo adquieren su rol causal, en virtud de sus propiedades formales (esto es, de su sintaxis).

Si tenemos las propiedades sintácticas y semánticas de una representación mental, tenemos todo lo que la constituiría como tal. Por las primeras, obtenemos la forma en que la representación se presenta (lo que, en algunas ocasiones, Fodor llama su “modo de presentación”); por las segundas, obtenemos su significado, expresado en su contenido. Dado esto, el carácter relacional de los estados intencionales puede considerarse como triádico: esto es, la relación entre un organismo, una estructura sintáctica y su contenido.

En una frase, creer-que-P para un organismo O equivale a que O tenga instanciada en su cerebro una muestra de un símbolo mental (dadas sus propiedades sintácticas) que significa que P (dadas sus propiedades semánticas). O, con palabras de Fodor: “creer que tal y cual es tener en la cabeza de una cierta manera una muestra de un símbolo mental que significa que tal y cual”<sup>45</sup>. De este modo, las propiedades sintácticas dan cuenta del rol causal de los estados mentales mientras que las propiedades semánticas dan cuenta de que un estado mental particular tenga el contenido que tiene, esto es, dan cuenta de su evaluación semántica<sup>46</sup>. Me concentraré ahora en cómo, exactamente, la teoría representacional da cuenta de esto último.

Como dije antes, T2 puede considerarse una versión de la tesis de la intencionalidad intrínseca. A pesar de que, dada la semántica composicional, el contenido de las representaciones mentales más complejas depende del contenido de sus representaciones individuales constituyentes, sigue en pie la pregunta acerca de cómo estas últimas adquieren sus propiedades semánticas. La respuesta es sencillamente que las propiedades semánticas de las representaciones mentales no son adquiridas. Y, más aún, la intencionalidad de las representaciones mentales es primitiva, irreducible a otras

---

<sup>45</sup> Fodor (1987, p. 38).

<sup>46</sup> Recuérdese que estas dos características son condiciones para una psicología de sentido común (p. 31), condiciones de las que Fodor puede dar cuenta con su noción de representación mental. Este punto es muy importante y será retomado más adelante.

entidades, propiedades o procesos y la fuente última de la intencionalidad de cualquier otra entidad, propiedad o proceso (o lo que fuera) que podamos considerar intencional.

De acuerdo con esto, Fodor entiende que hay, tanto a nivel ontológico como a nivel explicativo, una cadena de derivación de las propiedades semánticas de los objetos intencionales por la que la intencionalidad de las actitudes proposicionales es anterior a la intencionalidad de los lenguajes naturales y, a la vez, la intencionalidad de las representaciones mentales es anterior a la de las actitudes proposicionales. La intencionalidad de los niveles posteriores se reduce a la de los anteriores y cualquier instancia de intencionalidad intrínseca, no derivada, es propiedad exclusiva de las representaciones mentales en tanto entidades discretas del cerebro.

Como podrá parecer en una primera aproximación, esta postura es explicativamente muy “pesada” ya que conlleva una serie de consecuencias teóricas fuertes que necesitan de justificación y desarrollo. Voy a destacar aquí dos de ellas. En primer lugar, T2 lleva a incorporar un tipo de realismo que viene a sumarse al realismo acerca de las actitudes proposicionales que presenté antes (p. 33). No sólo la teoría representacional de la mente es realista acerca de los estados intencionales, sino que también lo es acerca de su contenido: esto es, acepta el *realismo intencional*, según el cual hay una cuestión de hecho acerca de cuál es exactamente el contenido de un estado intencional particular.

Nótese que este tipo de realismo es paralelo e independiente con respecto al realismo acerca de los estados intencionales. La idea novedosa aquí es que las propiedades semánticas (o intencionales) son reales y constitutivamente propias de las representaciones mentales. El carácter de “real” del contenido (esto es, de las propiedades intencionales) de las representaciones está en función de la caracterización de estas últimas como entidades discretas instanciadas físicamente en el cerebro: una de las propiedades inherentes a dichas entidades es su contenido intencional. De este modo, el realismo fodoriano se bifurca en un realismo acerca de los estados intencionales y un realismo acerca del contenido de dichos estados. Un “Realismo de peso pesado” (*industrial strength Realism*), diría Dennett (1991c).

Otra consecuencia importante que se sigue de T2 es que su aceptación nos deja inmediatamente con el llamado “problema de la naturalización del contenido” de los estados mentales, esto es, el problema de dar condiciones suficientes para que un estado tenga el contenido que tiene y, a la vez, hacerlo en términos estrictamente no intencionales. El problema surge inmediatamente si además se defiende, como lo hace

la gran mayoría de los filósofos de la mente, alguna versión de fisicalismo, esto es, la idea de que todo lo que hay es en última instancia físico. De acuerdo con esto, si lo intencional es una propiedad real de las cosas, sólo puede serlo en virtud de alguna relación *reductiva* (de identidad o superveniencia<sup>47</sup>, por ejemplo) con otras propiedades que no sean en sí mismas intencionales. Dado su fisicalismo, por una parte, y dado su realismo intencional, por otra, Fodor necesita adoptar una postura reduccionista acerca de las propiedades semánticas, proveyendo de este modo condiciones (especificadas en términos no intencionales y no semánticos) suficientes para la semántica del mentalés: es decir, suficientes para que un símbolo primitivo del mentalés tenga una cierta interpretación en un cierto contexto. La respuesta fodoriana a este problema constituye la cuarta tesis (T4) de la teoría representacional de la mente y me ocuparé de ella más adelante.

La tercera tesis (T3) de la teoría es la siguiente: “[e]l pensamiento es computación”<sup>48</sup>. Se suele subdividir la teoría fodoriana en la teoría representacional de la mente, que versaría sobre los estados mentales, y la teoría computacional de la mente, que versaría sobre los procesos mentales. El núcleo de esta última es T3. La idea básica es que los procesos mentales consisten en computaciones sobre representaciones mentales; lo que se entiende por computación aquí es la manipulación físicamente instanciada de símbolos (esto es, instancias de representaciones mentales concebidas como símbolos del lenguaje del pensamiento) en virtud de sus propiedades sintácticas o formales. Esta manipulación se da en términos de relaciones causales entre símbolos, que a la vez preservan sus propiedades semánticas. T3 proporcionaría de este modo un modelo para dar cuenta del realismo acerca de las actitudes proposicionales (p. 33) en tanto que los diferentes estados tipo de actitud proposicional (creencia, deseo y demás) podrían identificarse con diferentes conjuntos de relaciones computacionales entre representaciones con el contenido expresado por la actitud.

Otro aspecto importante de T3 es que las manipulaciones de símbolos operan no arbitrariamente sino de acuerdo con reglas, ya que de otro modo, por ejemplo, no podría explicarse el que una cadena de pensamiento preserve la verdad de sus constituyentes (las propiedades semánticas se preservan). Cabe aclarar que la teoría representacional de la mente no está comprometida con la representación explícita de estas reglas sino sólo

---

<sup>47</sup> Brevemente, se entiende que un conjunto de fenómenos, A, superviene sobre otro conjunto, B, si y sólo si dos objetos no pueden diferir respecto de A sin diferir también respecto de B.

<sup>48</sup> Fodor (1998, p. 27).

de los contenidos de los símbolos sobre los cuales aquellas operan. La posición que Fodor toma respecto de las reglas es que éstas podrían estar representadas explícitamente, pero que no es necesario que así sea. El punto importante es que haya algún mecanismo por medio del cual las propiedades semánticas de los símbolos sean respetadas de manera fiable por sus propiedades sintácticas, y la respuesta que da Fodor a esto está en su recurso a la metáfora del ordenador.

La idea detrás de la metáfora del ordenador, debida principalmente al matemático Alan Turing, es entender la mente básicamente como una especie de ordenador: fundamentalmente, una máquina de manipulación de información en términos de las propiedades sintácticas de los símbolos sobre los que opera. Como mencionaba antes, la sintaxis es considerada por Fodor como una propiedad física de orden superior: según esto, lo que conecta las propiedades semánticas inherentes a un símbolo con sus propiedades causales, esto es, susceptibles de ser manipuladas físicamente, es su sintaxis. El punto de la metáfora del ordenador es que ejemplifica un caso en que las relaciones semánticas entre símbolos son respetadas por las relaciones sintácticas correspondientes. De acuerdo con esto, los ordenadores “son precisamente entornos en los que la sintaxis de un símbolo determina su papel causal de una manera que respeta su contenido”<sup>49</sup>. La metáfora solucionaría el problema de la mediación entre propiedades causales y semánticas de los símbolos en tanto confirma la posibilidad de la co-ocurrencia fiable de ambas en un medio físico.

La última tesis<sup>50</sup> de la teoría representacional de la mente, T4, es la siguiente: “[e]l significado es información”<sup>51</sup>. T4 es parte de la solución fodoriana al problema de la naturalización del contenido que, como veíamos, se le plantea dado su fisicalismo y especialmente su defensa de una intencionalidad intrínseca. La teoría completa propuesta para hacer frente a dicho problema es conocida como la Teoría de la Dependencia Asimétrica (Fodor, 1990). Mi interés aquí no está tanto en los detalles de la teoría sino en determinar sus presupuestos y las razones para su postulación.

La teoría de la dependencia asimétrica se ubica en la tradición de las semánticas informacionales o basadas en la información (*information-based semantics*), cuyo precursor directo es el trabajo de Fred Dretske (1981). La idea básica que vertebra las semánticas informacionales es la de que el contenido mental es una especie de

---

<sup>49</sup> Fodor (1987, p. 41).

<sup>50</sup> En realidad Fodor agrega una quinta tesis respecto de los conceptos coextensivos, que no trataré ya que no es pertinente para la discusión.

<sup>51</sup> Fodor (1998, p. 30).

contenido informacional; a la vez, la noción de información que se asume se asienta en las correlaciones causales / nomológicas entre dos fenómenos, de los cuales se dice que uno lleva información sobre el otro. Dado este marco, el problema que la teoría de la dependencia asimétrica pretende enfrentar es el de determinar en qué consiste la posesión de un contenido intencional y qué hace que un estado intencional particular tenga el contenido determinado que lo identifica como tal.

Como ya había mencionado, esta tarea es asumida bajo la restricción de que la explicación proporcionada sea reductiva y, en este sentido, sea especificada en términos no intencionales o semánticos. Para este propósito, la noción dretskeana de información puede cumplir el rol requerido en tanto es objetiva y no apela a proceso alguno de interpretación por parte de un agente cognitivo externo: se asienta en la conexión nomológica que es posible inferir entre un fenómeno, esto es, la instanciación de una representación mental, y aquello que lo causa de manera fiable, esto es, alguna instancia en el mundo, conexión en virtud de la cual puede decirse que el primer fenómeno lleva información acerca del segundo (o mejor, en el caso de Fodor, lo representa).

Ahora, Fodor no identifica íntegramente el contenido mental con la información así entendida, y esto porque puede darse el caso de que una representación mental con contenido intencional sea el efecto fiable de cosas que no están en su extensión. Este problema es conocido como el problema de la disyunción, y a él debe enfrentarse todo aquel que pretenda dar cuenta de la determinación exhaustiva del contenido mental. Aceptar este problema implica también, de acuerdo con el planteo fodoriano, rechazar una explicación funcionalista acerca del contenido mental, en especial, la de las llamadas “semánticas del rol funcional” (*functional role semantics*): apelar a la noción de función lleva a la indeterminación del contenido (en términos de Fodor, al holismo del contenido) dada la indeterminación implícita que cualquier función conlleva en cuanto tal<sup>52</sup>.

La teoría de la dependencia asimétrica resuelve el problema de la disyunción apelando a una diferencia jerárquica entre tipos de relaciones informacionales. La idea es que algunas de estas relaciones son más básicas o fundamentales para la determinación del contenido: de acuerdo con esto, algunas relaciones sólo pueden sostenerse fiablemente en base a que se haya previamente establecido la relación

---

<sup>52</sup> El concepto de función, aplicado a la cuestión del contenido, es inherentemente holista dado que el que algo tenga o asuma una función depende de modo necesario de las relaciones que mantenga con otros objetos de su entorno y, en el caso de un sistema complejo, como un organismo, con otras funciones de sub-sistemas de éste.

fundamental. Esta última conecta nomológicamente la propiedad en el mundo con la representación instanciada cuyo contenido expresa dicha propiedad. Dicho de otro modo, la conexión que determina el contenido de una representación a partir de una propiedad en el mundo que no expresa dicho contenido es asimétricamente dependiente de la conexión nomológica fundamental entre una propiedad en el mundo y la representación cuyo contenido expresa la misma. Si esta conexión se quiebra, la primera también lo haría, pero no sucede a la inversa.

Otro punto a destacar es que no se requiere que la conexión nomológica para la determinación del contenido sea causal, es decir, que se establezca efectivamente, sino solamente que obedezca leyes contrafácticas por las cuales, si se dieran determinadas condiciones (la instanciación de una propiedad en el mundo), se seguirían determinados efectos (la instanciación de la representación mental que expresa dicha propiedad). La idea de Fodor es que podría haber leyes entre propiedades aún cuando éstas no puedan instanciarse (por ejemplo, aunque no existen unicornios, podría aún así haber una ley por la cual los unicornios causarían representaciones cuyo contenido expresara la propiedad de ser-un-unicornio).

Me detengo aquí en la exposición de la teoría representacional de la mente. He revisado los puntos más importantes de la teoría con un enfoque más bien metodológico y meta-teórico. Quiero recalcar que mi interés en la teoría fodoriana, como para el caso de la teoría dennettiana, se ubica en un plano argumentativo que apunta a la determinación del marco general para su defensa, la identificación de sus supuestos, la delimitación de las discusiones puntuales que abarca y la resultante concepción de la realidad mental que favorece. Estos puntos serán discutidos desde una perspectiva comparativa con la teoría de los sistemas intencionales en el capítulo 4. Me ocuparé ahora de justificar la inserción de la teoría representacional de la mente, tal como la he expuesto y en vista de dicho plano argumentativo, en lo que llamo “*enfoque anti-evolucionista acerca del contenido mental*”.

## ***2. Enfoque Anti-Evolucionista acerca del Contenido Mental***

En primer lugar, la teoría representacional de la mente podría inscribirse dentro de un enfoque anti-evolucionista dado que en ninguno de sus puntos se apela, ni siquiera de modo débil, a consideraciones evolucionistas. Aunque el caso de Fodor no es único ni

mucho menos, parece un punto importante para destacar el que, para la totalidad de los distintos temas que la teoría representacional abarca, desde diversas perspectivas y frente a diferentes críticas, no se considere pertinente apelar a la teoría evolucionista, ni en general ni recurriendo a alguna idea particular. Por una parte, como vimos, no hay ninguna tesis sustantiva de la teoría que incluya alguna referencia a la teoría de Darwin y, por otra parte, los argumentos a los que se apela para la defensa de dichas tesis tampoco recurren a ella. Quisiera decir algo sobre estos argumentos.

Fodor suele distinguir los argumentos para su teoría entre empíricos y no empíricos. Algunos de estos últimos serán tratados en el capítulo 4, pero vale anticipar aquí que todos ellos dependen de condiciones puestas o supuestas por la teoría y que son estrictamente independientes de un compromiso con la teoría darwinista. Por otra parte, hay sólo un argumento general que Fodor (1975, 1987) considera empírico. En algún grado, todas las tesis que constituyen la teoría representacional de la mente son defendidas apelando al uso implícito que de ellas harían la psicología y otras disciplinas cognitivas: más precisamente, se sostiene que dichas disciplinas asumen las tesis, esto es, las toman como supuestos aceptados de sus investigaciones. En particular, Fodor sostiene que hay por parte de estas disciplinas un compromiso ontológico con la noción fuerte de representación mental, tal como la presenta su teoría.

Ahora, más que un argumento empírico, éste pareciera ser un argumento hacia la mejor explicación en tanto que la idea básica es que, si nuestras mejores teorías actuales sobre el funcionamiento de los procesos cognitivos suponen alguna versión de la teoría, entonces la verdad de aquellas teorías apoyaría la verdad de ésta. Sea el argumento del tipo que sea, ciertamente no es empírico. O si lo fuera, lo sería solamente de modo muy indirecto ya que el argumento da por supuesta la adecuación empírica de un conjunto entre otros de teorías psicológicas actuales. Lo que me parece importante del argumento es que, además de subrayar el compromiso fodoriano con la investigación psicológica empírica, expresa el supuesto de que hay una conexión fuerte entre este tipo de investigación y la propuesta filosófica de Fodor: esto es, la idea de que la adecuación de la teoría representacional de la mente depende en una primera instancia de la adecuación de teorías psicológicas empíricas coherentes con ella. Mi discusión de la teoría en el capítulo 4 apuntará especialmente a rechazar este supuesto.

La falta de una mención explícita de la teoría evolucionista no es suficiente para juzgar las ideas fodorianas estrictamente como anti-evolucionistas. Sin embargo, Fodor expone en algunos lugares una crítica directa a la pertinencia de un recurso a dicha

teoría para el estudio filosófico sobre la mente. Estas escasas pero claras manifestaciones son suficientes para convertirlo en un defensor explícito del enfoque anti-evolucionista. La idea básica de Fodor (1994, pp. 19-20) es que el tipo adecuado de explicación psicológica es *sincrónico*; es decir, invoca entidades, procesos y propiedades que son operantes en el momento presente. Por otro lado, las explicaciones evolucionistas son históricas y apelan al proceso que llevó a la conformación de lo que sea que se esté explicando. En este sentido, no tienen utilidad para un estudio psicológico científico: “[l]a evolución quizás explica porqué hay más cosas que funcionan que cosas que no funcionan. Pero no explica *cómo* las cosas funcionan...”<sup>53</sup>.

Se encuentran otras explícitas declaraciones anti-evolucionistas de Fodor, y puntualmente en contra de Dennett, en Dahlbom (1993, pp. 74-77). Ahí Fodor expone una crítica a lo que llama el “argumento evolucionista de Dennett para el principio de verdad” (a pesar de que, estrictamente, no existe un tal argumento). El principio de verdad consiste en suponer que las adscripciones intencionales representan las creencias de un organismo como mayormente verdaderas, y lo que el (supuesto) argumento de Dennett, *via* la apelación a Darwin (y, en particular, al carácter adaptativo de la verdad de las creencias), pretende mostrar es que es necesario que esto sea así: es decir, concebir un organismo como creyente de verdades es constitutivo de la adscripción de contenidos intencionales como tal. De este modo, toda atribución de contenido depende de la particular interpretación de lo que las creencias verdaderas (y los deseos apropiados) deben ser para un organismo, dada su situación. Fodor hace nuevamente hincapié en el carácter histórico-causal de la teoría evolucionista para argumentar que no puede apelarse a una teoría de este tipo para sostener que la adscripción de contenidos es, en el fondo, una cuestión de interpretación:

Es, en breve, extraño cómo Dennett piensa que recurrir a la teoría darwinista –que es, después de todo, una historia causal sobre los *mecanismos de especiación*– pueda revelar un “elemento de interpretación” en la adscripción de contenido<sup>54, 55</sup>.

La explicitada limitación en el papel explicativo de la teoría evolucionista muestra otro aspecto del anti-evolucionismo fodoriano. No sólo hay, por parte de Fodor, un rechazo a un uso directo de la teoría evolucionista por su aducida inutilidad para la filosofía de la mente, sino que tampoco se reconoce la posibilidad de hacer un uso, por

---

<sup>53</sup> Fodor (1994, p. 20). La traducción es mía.

<sup>54</sup> En Dahlbom (1993, p. 74). La traducción es mía.

<sup>55</sup> Retomaré esta crítica fodoriana en el apartado 4.3 (p.103).

así decir, *indirecto* de ella. Es decir, por una parte, se niega la pertinencia de argumentos que apelen directamente a la teoría pero, por otra parte, tampoco se reconocen modos alternativos justificados en que ésta pueda jugar un rol en el estudio de la mente. Como vimos en el caso de la teoría de los sistemas intencionales, pueden distinguirse dos maneras en que el evolucionismo puede adoptarse: lo que, para el caso de Dennett, llamaba carácter gradualista metodológico como distinto del carácter gradualista sustantivo (p. 13). Esta distinción puede hacerse en la medida que se acepte que la teoría evolucionista puede ofrecer un *modelo* explicativo determinado, que ponga restricciones meta-teóricas a la teoría de la mente que se esté elaborando. En mi opinión, este tipo de aporte es uno de los que la reflexión filosófica puede beneficiarse mayormente<sup>56</sup>.

De acuerdo con esto, quiero aclarar nuevamente que la identificación de un modelo evolucionista para la teoría de Dennett como contrapuesto a un modelo anti-evolucionista para la de Fodor apunta sobre todo a la inserción de las tesis particulares de ambas teorías en un marco general y flexible, teórico y metodológico, que proporciona un trasfondo para la resolución de los problemas particulares tratados por ellas. Dicho de otra manera, en general, el carácter evolucionista o anti-evolucionista de la teoría no afecta directamente las tesis particulares que se defienden (aunque, como vimos, en el caso de Dennett se recurre en diferentes puntos a argumentos que apelan directamente a la teoría evolucionista). Por otro lado, los supuestos de la teoría, el alcance de sus tesis, los objetivos que pretende satisfacer y el estilo argumentativo que presenta son todos puntos afectados por la relación indirecta que una teoría filosófica particular tiene o deja de tener con el planteo evolucionista.

Para concluir este capítulo, e introducir algo de las discusiones por venir, quisiera rápidamente mostrar el grado en que la misma interpretación de Fodor de las ideas evolucionistas es cuestionable. En su reseña de *Darwin's Dangerous Idea*, Fodor toma la posición de Gould en su crítica al adaptacionismo. Cito a Dennett, citando a Fodor: “[h]abría que esperar y ver cómo, y *si*, nuestras mentes evolucionaron. Hasta el momento, *no tenemos los datos*”. Y:

De todos modos, lo que no está disponible es el rumbo en el que Dennett parece estar embarcado: no hubo diseñador, pero aún así el reloj fue diseñado. *Eso simplemente no tiene sentido*<sup>57</sup>.

---

<sup>56</sup> En el apartado 4.4 y en la Conclusión haré algunas reflexiones sobre la influencia de los enfoques evolucionista y anti-evolucionista en la relación que una teoría filosófica de la mente puede tener con otras disciplinas, así como en la misma concepción de la práctica filosófica y sus objetivos.

<sup>57</sup> Dennett (1996b, pp. 264-265). La traducción y el subrayado son míos.

A la luz de mi análisis del evolucionismo en la teoría de los sistemas intencionales, estas (sorprendentes) afirmaciones fodorianas dan una idea de cómo ambos autores no discrepan “sin más” sobre el papel de la teoría evolucionista para la reflexión filosófica sobre la mente, sino que esta discrepancia radica en parte en su interpretación de la teoría darwiniana (incluso, como las citas revelan, más allá de la disputa adaptacionista) y de la importancia teórica y meta-teórica que, dada dicha interpretación, la teoría puede tener.

Doy por concluida la presentación de la teoría fodoriana. El carácter anti-evolucionista general y de las tesis particulares de la teoría será explicitado en detalle como parte de la exposición de las críticas dennettianas a Fodor. Baste por ahora subrayar la falta del recurso explícito a consideraciones evolucionistas (el “no-evolucionismo” fodoriano), la crítica fodoriana a la pertinencia de dicho recurso y la posibilidad de atribuir a la teoría representacional de la mente un marco metodológico que entraría en conflicto con el modelo evolucionista asumido por Dennett. Que este marco puede atribuirse a la teoría fodoriana es una de las tesis que voy a explicitar y defender a lo largo de mi discusión comparativa entre ambas posiciones. Ahora paso al análisis de una distinción, a mi parecer, fundamental en el marco del pensamiento dennettiano, que vertebrará la disputa Dennett / Fodor a la luz de los enfoques evolucionista y anti-evolucionista acerca del contenido mental: la distinción dennettiana entre “creencia” y “opinión”.

---

---

## La Distinción entre Creencia y Opinión

### *1. Presentación y Análisis Preliminar de la Distinción*

La distinción técnica entre “creencias” y “opiniones” fue primeramente propuesta por Dennett en el artículo “How to Change Your Mind”<sup>58</sup> de 1977, como resultado de la elaboración de una distinción previa, propuesta por Ronald de Sousa (1971), entre lo que este último llama “creencia”, por un lado, y “asentimiento”, por otro. A pesar de que Dennett retoma la distinción en varios trabajos posteriores, no la elabora con detenimiento y sólo la emplea como una herramienta útil en el contexto de diversos debates. No hay, aparte del artículo mencionado, ningún otro lugar en la obra dennettiana en que la distinción reciba ulteriores especificaciones<sup>59</sup>. A la vez, la distinción no cumple, en ninguna de las discusiones para las que sería pertinente, un rol particularmente importante, sino que su relevancia e implicaciones son muchas veces dadas por supuesto.

Creo que esta falta de elaboración es criticable<sup>60</sup> en tanto que voy a sostener que la distinción es un punto fundamental, tanto sustantivo como metodológico, para la teoría del contenido de Dennett y, especialmente, respecto del modo en que se enfrenta al conjunto de teorías que se le oponen (puntos que desarrollo en el apartado 3.3). Este capítulo quiere hacer frente a esta situación, estudiando los diferentes aspectos de la distinción y cómo pueden desarrollarse, a la vez tratando de enmarcarla dentro del enfoque evolucionista dennettiano y evaluando sus implicaciones para la filosofía de la mente. Estas implicaciones serán la base para la discusión de los debates claves entre la teoría de los sistemas intencionales y la teoría representacional de la mente, discusión a la que dedicaré el próximo capítulo. Dada la mencionada infra-elaboración del tema,

---

<sup>58</sup> Dennett (1978a).

<sup>59</sup> Quizás exceptuando el caso de Dennett (1991d), que sin embargo tampoco ofrece un desarrollo sustantivo.

<sup>60</sup> En esto me acompañan tanto Andy Clark (1990) como Keith Frankish (1998): el primero habla del “tratamiento sub-desarrollado de la distinción” mientras el segundo critica la falta de una elaboración sistemática de sus términos. También Bo Dahlbom (1993, p. 162) y el mismo Dennett (Op. Cit., p. 229) mencionan el asunto.

tampoco ha habido en la literatura mayor dedicación a él; por esto, mis únicos puntos de apoyo bibliográfico, fuera de la obra de Dennett, serán los artículos que Andy Clark (1990) y Keith Frankish (1998) dedican específicamente a la distinción.

Es pertinente aclarar, de modo preliminar, que la distinción bajo estudio funciona netamente en el nivel personal de explicación (ver p. 9): la distinción no tiene repercusión directa alguna a nivel sub-personal y esto, a la vez, responde al hecho de que ambas, creencias y opiniones, son atribuidas desde la perspectiva de la estrategia intencional. Concretamente, hay una diferencia específica y crucial que justifica teóricamente la distinción: esta diferencia está dada por el componente *lingüístico* presente en la conformación del estado atribuido en el caso de las opiniones y ausente, en el caso de las creencias.

En este sentido, las creencias estarían definidas como estados mentales *básicos*, disposicionales, no ocurrentes e indeterminados, atribuibles desde la estrategia intencional<sup>61</sup>; las opiniones, por otra parte, serían estados *derivados*, lingüísticamente “infectados”, y, en tanto tales, radicalmente más específicos y con valores de verdad determinados. La distinción, entonces, sirve en una primera instancia para marcar una diferencia entre aquellos organismos capaces de manipular un lenguaje, por un lado, y el resto de los sistemas intencionales, por otro, distinción que, como veremos, viene en definitiva a circunscribir el grupo de los sistemas intencionales humanos frente al resto. La distinción también sirve para diferenciar, en las personas, distintos tipos de estados intencionales, aunque este punto es problemático.

Es importante tener en cuenta que, para el caso particular concreto de nuestras atribuciones mentalistas paradigmáticas, esto es, referidas a personas adultas lingüísticamente competentes, los dos términos de la distinción no pueden deslindarse claramente. En cada caso particular, no puede dirimirse de forma estricta si se está frente a una creencia o, alternativamente, a una opinión porque, como en seguida veremos, ambas son manifestaciones diferentes de la misma realidad y, como tales, siempre que intervenga el elemento lingüístico, no pueden considerarse por separado<sup>62</sup>.

---

<sup>61</sup> Es interesante notar el parecido que esta noción de creencia tiene con la noción wittgensteineana en *Über Gewissheit* (1969): las creencias como estados no concientes en los que se apoyan los pensamientos concientes (el concepto de opinión también es usado por Wittgenstein, aunque en escasas oportunidades, en *Philosophische Untersuchungen* (1953), y en un sentido similar al empleo que hace Dennett).

<sup>62</sup> Estas consideraciones son concretamente aplicables a los sistemas intencionales humanos, una vez que adquieren competencia lingüística, dado que, entre otras razones, es siempre posible para una creencia convertirse en un estado ocurrente para el agente.

En relación con esto, creo que sería ir en contra de las intenciones de Dennett tomar los términos de la distinción como, si se me permite la expresión, “referencialmente nítidos”; esto es, utilizar los términos como identificando casos claros de tipos diferentes de fenómenos (casos de creencia y casos de opinión). Me concentraré más adelante en los diferentes motivos que hay detrás de la propuesta del par creencia / opinión, pero cabe destacar aquí que dicha propuesta no pretende individuar dos fenómenos absolutamente distinguibles y, mucho menos, independientes uno del otro. Más bien, se quieren destacar ciertos rasgos característicos de nuestras atribuciones mentalistas (propiaamente humanas) y del conjunto de fenómenos a los que refieren.

A pesar de las aclaraciones anteriores, que más bien van en contra de un uso inapropiado de los términos, las características definatorias de los mismos los distinguen radicalmente y están sobre todo relacionadas con la diferencia crucial que antes destacaba: la presencia o ausencia del elemento lingüístico. El único factor en común que puedo reconocer entre los términos es el hecho de que ambos son atribuibles desde la perspectiva que asumimos al adoptar la actitud intencional. Como vimos en el capítulo 1, el único criterio para el caso de las creencias es su susceptibilidad de ser atribuibles predictivamente desde la adopción de esta estrategia; para el caso de las opiniones también adoptamos la actitud intencional, pero los mecanismos específicos de atribución funcionan, como veremos, de modos diametralmente distintos. Para entender estas diferencias, deberemos concentrarnos en (algunas de) las innovaciones específicas que el uso del lenguaje natural brinda a nuestras capacidades cognitivas.

En el artículo “Learning and Labeling” (1993), Dennett, en respuesta a la crítica de Clark y Karmiloff-Smith en “The Cognizer’s Innards” (1993a), se concentra en el carácter estructurado y simplificador de nuestro lenguaje natural como un factor determinante para la amplificación del poder redescriptivo de nuestras capacidades cognitivas. La diferencia sustancial que introduce el lenguaje, y que justifica teóricamente la distinción creencia / opinión, se debe a algunas de sus características peculiares, en especial, su carácter estructurado, combinatorio y recursivo, así como también al carácter fijo y simplificador de sus términos.

Todas estas características justifican la restricción excluyente del término “opinión” para los casos de atribuciones a seres humanos lingüísticamente competentes. Los casos de proto-lenguajes en comunidades animales o los mismos balbuceos intencionados en infantes humanos están descartados justamente porque no incorporan las mencionadas características propias de nuestro lenguaje natural. Como veremos en

el siguiente apartado, estas mismas características consideradas en conjunto reflejarán una situación que acentuará fuertemente el carácter excluyente de las opiniones.

Respecto del carácter fijo y simplificador de los términos de los lenguajes naturales, Dennett subraya el rol de la práctica de poner etiquetas o “etiquetamiento” (*labeling*) a experiencias particulares y conjuntos de experiencias como un modo de crear un conjunto de objetos fijos y, por tanto, fácilmente manipulables:

Una vez que hemos creado etiquetas, y el hábito de “pegarlas” [*attaching them*] a circunstancias experimentadas, hemos creado una *nueva clase de objetos* que pueden a la vez volverse los objetos de toda la maquinaria de reconocimiento de patrones, de construcción de asociaciones, y demás<sup>63</sup>.

La ventaja central de esta creación de objetos es que da lugar a una tarea perceptiva simple (esto es, el reconocimiento de las etiquetas) para la subsiguiente operación sobre características del complejo medio ambiental, que de este modo se vuelven salientes y concretas.

Es principalmente por este motivo que el mecanismo de asociación de los términos lingüísticos puede a su vez cobrar un carácter fuertemente combinatorio y recursivo (el segundo conjunto de características destacadas por Dennett), dando así lugar a una arquitectura conceptual que puede asumir una altísima complejidad estructural. Estas características en conjunto, luego, tienden a favorecer una concepción del lenguaje como una suerte de *modalidad perceptiva* y, a la vez, una modalidad perceptiva que es novedosa en su funcionamiento. Baste este esbozo, por ahora, respecto del cambio drástico que produce el lenguaje para la conformación de las opiniones, tema que será retomado en el siguiente apartado para discutir el gradualismo de la distinción.

Hago ahora una digresión para concentrarme en algunos aspectos cruciales de la interpretación de la distinción creencia / opinión. Creo que puede advertirse una tensión en la escueta caracterización que hace Dennett de las opiniones; se las podría interpretar de dos maneras diferentes: de un modo *restringido* y de un modo *extendido*. En el primer caso, las opiniones vendrían a ser sencillamente, como el autor repite en algunas ocasiones, “apuestas sobre la verdad de oraciones en un lenguaje que entendemos”<sup>64</sup>. En este caso, las opiniones serían entendidas de modo netamente “intelectualista”, es decir, como *eventos* particulares de asentimiento conciente ante una oración determinada,

---

<sup>63</sup> Dennett (1993, p. 543). La traducción y el subrayado son míos.

<sup>64</sup> Dennett (1998a, p. 89). La traducción es mía.

como actos aislados de decisión dentro de un contexto epistémico<sup>65</sup>. En esta versión, mis reparos anteriores (pp. 46-47) contra una interpretación que tome a los términos de la distinción (en este caso, las opiniones) como fenómenos independientes estarían fuera de lugar. Podría de hecho identificarse el acto-de-tomar-por-verdadero con el fenómeno “opinión”, distinguible e independiente del fenómeno “creencia”<sup>66</sup>.

Por otra parte, en lo que sería el modo extendido de interpretación, el concepto es usado en repetidas ocasiones con el objetivo general de marcar la diferencia sustancial que la incorporación del lenguaje en la vida mental de los seres humanos tendría sobre las creencias. En este sentido, las opiniones vendrían a ser una cosa misma con las creencias, esto es, *estados* (y no eventos) atribuibles desde la actitud intencional, una vez que sufren la influencia distorsionante del lenguaje. Entiendo esta ambigüedad como una consecuencia del ya mencionado sub-desarrollo de la distinción; digo esto específicamente porque, a pesar de que estoy defendiendo la segunda interpretación de la distinción, creo que hay elementos de la interpretación restringida que influyen en la conformación de las opiniones como estados-parecidos-a-creencias<sup>67, 68</sup>.

Mi punto es el siguiente: la idea de las opiniones como apuestas sobre la verdad de enunciados determinados (esto es, el núcleo de la interpretación restringida) debe interpretarse como el eje que proporciona el vínculo entre lo que antes consideraríamos una creencia y la opinión resultante. Lo que sigue a continuación debe entenderse como una abstracción de un proceso general y difuso, y no como la descripción detallada de un proceso concreto. Me refiero al proceso de transformación de la creencia en opinión<sup>69</sup>.

Dennett retoma la idea de de Sousa según la cual habría algo así como dos niveles distintos funcionando en el caso de la creencia humana: tendríamos como base

---

<sup>65</sup> Esta interpretación haría de la distinción dennettiana una continuación muy fiel de la distinción de de Sousa entre creencias y asentimientos. Una expresión clara de este tipo de interpretación la da Frankish (1998, p. 430): “[l]as opiniones, según parece, son solamente oraciones que uno escoge y está dispuesto a tener como verdaderas” (la traducción es mía). Dada esta interpretación, Frankish pasa luego a proponer lo que llama un “rol cognitivo” para las opiniones.

<sup>66</sup> Estrictamente hablando, sin embargo, esto valdría solamente para las opiniones así entendidas, ya que las creencias, tal como se configuran desde la óptica de la teoría de los sistemas intencionales, no son *fenómenos* en sentido fuerte (ver apartado 1.1).

<sup>67</sup> Además del hecho de que las varias apelaciones de Dennett a la distinción entre creencias y opiniones, posteriormente a su presentación en “How to Change Your Mind”, se encuadran en un contexto que requiere un rol para el concepto de “opinión” mayor al que le otorgaría la interpretación restringida. Ampliaré sobre esto en el apartado 3.3.

<sup>68</sup> Un punto importante a tener en cuenta es que ambas interpretaciones no son contradictorias o incompatibles, sino que refieren a aspectos diferentes de lo mismo.

<sup>69</sup> Esto no tiene que implicar necesariamente que las opiniones sólo puedan surgir como transformaciones de creencias pre-existentes.

nuestras creencias y deseos del tipo que compartimos típicamente con los animales, vagamente, estados difusos, disposicionales, no ocurrentes y no proposicionales, y, a partir de esta base, realizaríamos “saltos de asentimiento” (con palabras de Dennett) para la conformación de las opiniones. La razón de que la operación hacia el segundo nivel de creencia sea considerada un “salto” se debe al carácter fijo de los términos lingüísticos, que arriba destacaba.

El asentimiento que damos (necesariamente) por medios lingüísticos conlleva el recorte y la simplificación del entorno complejo en el que, dada nuestra interacción con él, nuestras creencias y deseos se generan y al que nuestras opiniones hacen referencia. Como Dennett destaca, la relación entre, por un lado, creencias y deseos y, por otro lado, el comportamiento o la conducta que éstas regulan es absolutamente directa:

El propio comportamiento es consonante con las propias creencias [término técnico] “automáticamente”, ya que así es cómo a fin de cuentas individuamos creencias y acciones<sup>70</sup>.

Por otra parte, dado el inevitable salto lingüístico a través del cual las opiniones se configuran, la relación entre éstas y la conducta será siempre irreductiblemente indirecta.

Ahora bien, hay un factor central a tener en consideración: esto es, el hecho de que, literalmente, vivimos embebidos de lenguaje. Constantemente estamos rodeados de palabras y conceptos, sean éstos oídos, recordados, dichos, imaginados, escritos o leídos. Este es un hecho trivial pero con una enorme importancia en este punto: es el hecho que justifica el paso de la interpretación restringida de las opiniones a la extendida. Nuestra incesante gimnasia mental por medios lingüísticos tiene inevitablemente un efecto profundo sobre nuestra vida mental que, como tal, tiende a redimensionarse por completo.

Más allá de la importancia que pueda tener la acción puntual de asentir, de apostar por la verdad de una oración, es el efecto que tal actividad tiene sobre nosotros el punto fundamental de la distinción creencia / opinión. En este sentido, lo central no es tanto el acto de asentimiento como tal, esto es, la resolución a la que llegamos, sino el continuo proceso que lleva hasta ella: es decir, el hecho de que estemos constantemente considerando el interés, el valor, la utilidad y, en general, las implicaciones de las

---

<sup>70</sup> Dennett (1978a, p. 307). La traducción es mía.

oraciones que, del modo que sea, llegan a nosotros. Este hecho a su vez genera el fenómeno novedoso, paralelo a las opiniones, del *pensamiento conciente*<sup>71</sup>.

Hay que recordar aquí que esta capa superior de incesante consideración de ideas en soledad o con otras personas nunca es estrictamente distinguible del estrato disposicional inferior, ni las operaciones de aquella pueden considerarse con total independencia de la influencia de éste<sup>72</sup>. Nuestra inmersión en el océano cultural de las palabras, desde nuestra llegada al mundo en adelante, hace imposible una vivisección nítida de los elementos estrictamente disposicionales respecto de aquellos estrictamente culturales de nuestra vida mental. Los dos coexisten y se entremezclan ante nuestra total inconciencia. Lo único que puede estudiarse es el resultado complejo de esta interacción, y la distinción entre creencias y opiniones justamente pretende destacar algunas características de este fenómeno complejo, con el objetivo principal de dar cuenta de la alteración que el lenguaje proporciona sobre nuestra mentalidad y, con esto, de la diferencia entre los estados intencionales en general respecto de los estados intencionales en el caso humano.

Ahora bien, es central entender el carácter fuertemente *derivado* que asumen las opiniones dentro del esquema de Dennett. Esto es, las opiniones, entendidas como estados de creencia infectados lingüísticamente, son un producto especializado que atraviesa un proceso muy importante de transformación, que deja poco lugar para la analogía con las meras creencias. Me parece que un punto de dificultad en la interpretación de la distinción en cuestión está en conciliar dos características aparentemente encontradas: por un lado, el hecho de que las opiniones son, en un sentido importante, un desarrollo de los estados de creencia una vez sometidos a un proceso de elaboración por medios lingüísticos; por otro lado, el hecho de que la conexión que puede trazarse comparativamente entre ambos estados es, como mínimo, fuertemente indirecta, y esto en tanto que la incorporación del lenguaje a nuestra vida mental humana produce una verdadera revolución cognitiva al incorporar mecanismos y consecuencias concretas radicalmente diferentes de los que podemos atribuir para dar cuenta del comportamiento de cualquier organismo no lingüístico. Este aparente contraste puede reformularse, como a mi entender ya lo hizo Clark, como una falla en el

---

<sup>71</sup> En este sentido debe entenderse la afirmación de Dennett en “Making Tools for Thinking” (2000, p. 20) de que los chimpancés, nuestros parientes evolutivos más cercanos, no deben considerarse capaces de pensamiento.

<sup>72</sup> Esto hace particularmente apremiante la interpretación extendida de la distinción, aunque, como indiqué en la nota 11, aceptarla no implica necesariamente negar la viabilidad de la interpretación restringida.

aspecto gradualista de la postura dennettiana (y, forzosamente, en su evolucionismo). En este punto me voy a concentrar en el siguiente apartado.

## ***2. Carácter Gradualista de la Distinción***

Aunque todavía no me he dedicado al rol (o, como veremos, los roles) que la distinción entre creencias y opiniones juega dentro del contexto más amplio de la teoría de los sistemas intencionales y de su debate frente a teorías rivales, la distinción, tal como la he presentado hasta aquí, parecería marcar un quiebre conceptual fuerte, sobre todo si se la interpreta en el marco de la tendencia gradualista general de la teoría dennettiana<sup>73</sup>. La razón de esto está en que la distinción de hecho se hace eco de una diferencia absolutamente tajante, esto es, la capacidad para la manipulación de un lenguaje articulado o, más específicamente, cualquiera de los lenguajes propios de las distintas comunidades humanas.

Dejando fuera de consideración datos poco relevantes como una mayor o menor complejidad propia del lenguaje en cuestión o un mayor o menor nivel de competencia por parte del usuario, puede considerarse que estamos frente a una distinción fuerte, sin casos intermedios, por así decir. Más aún, Dennett considera que esta capacidad exclusivamente humana comporta un avance cognitivo tan importante que pone al ser humano en tanto especie “directamente *fuera* de la escala de inteligencia que ubica al cerdo sobre la lagartija y la hormiga sobre la ostra”<sup>74</sup>. Específicamente estos puntos fueron tomados por Andy Clark, quien, desde la publicación de su “Belief, Opinion and Consciousness” en 1990 y en varias publicaciones sucesivas, viene protagonizando junto con Dennett una disputa respecto de la distinción creencia / opinión (entre otros puntos), en el marco del problema de determinar el papel que el lenguaje natural ocuparía en la conformación de nuestra particular (característicamente humana) vida mental.

Antes de pasar puntualmente a Clark, quiero explorar algunas ideas de la teoría dennettiana que pueden dar pie para una interpretación gradualista de la distinción creencia / opinión. El modelo evolutivo de la torre de la generación y la prueba (ver pp. 20-21) propone una serie de estructuras cognitivas simplificadas cuyo estadio más

---

<sup>73</sup> Ver apartado 1.2.

<sup>74</sup> Dennett (1994, p. 164). La traducción y el subrayado son míos.

avanzado es el de las criaturas gregorianas, caracterizadas por su capacidad para diseñar y utilizar herramientas para facilitar sus tareas cognitivas. Dennett aquí se hace eco de la idea de Richard Gregory (1981) según la cual la posesión y el manejo de una herramienta no sólo es indicio de inteligencia por parte del usuario capacitado, sino que es un medio que en un sentido muy directo le otorga inteligencia y, de este modo, amplifica su capacidad para disponer de los recursos ambientales a su alcance.

Estas herramientas de las criaturas gregorianas, entendidas como medios diseñados para algún objetivo específico, constituyen el modo que éstas tienen para delegar su carga cognitiva en el medio ambiental, haciendo así su tarea más fácil y con un potencial aumentado. Mientras más información suponga la fabricación de la herramienta, más inteligencia podrá otorgar a su usuario, y las palabras, las “herramientas mentales” en términos de Gregory, son el ejemplo paradigmático de herramientas que condensan máximamente información acumulada en su manufactura.

Ahora, para que una herramienta pueda ser utilizada debe ser manipulable de algún modo y, por esto, tiene que ser concreta y fácilmente distinguible. La operación sobre conceptos e ideas que, como vimos, constituye el estrato superior de las opiniones y del pensamiento conciente justamente parecería no ser manipulable en este sentido directo y concreto. Pero, por otra parte, éste es exactamente el sentido por el que el lenguaje, de acuerdo con mi interpretación de Dennett, podía verse como una nueva modalidad perceptiva, dadas algunas de sus características principales (ver p. 48). En este sentido, las palabras, como elementos fijos, distinguibles y re-identificables, son entendidas como *manipulables* en un sentido básico que es esencialmente el mismo que se aplica a cualquier herramienta externa concreta.

Estas consideraciones pueden hacerse extensivas al plano del pensamiento abstracto propiamente dicho, si además aceptamos la idea de las palabras como *predecesoras* de los conceptos, idea que Dennett propone en “Learning and Labeling”. El hábito de etiquetar experiencias en los niños que están desarrollando su capacidad lingüística es el proceso de conformación de un nuevo conjunto de objetos que el niño aprende a distinguir y aplicar en situaciones-tipo, ante los estímulos y las correcciones de quienes lo rodean. La práctica inicial de constante “manoseo” irreflexivo de las palabras pronunciadas en el entorno del niño se va sofisticando y articulando hasta alcanzar el nivel adulto de manejo conciente de conceptos e ideas. El punto es que los conceptos que intervienen en la formulación e intercambio de las opiniones pueden cumplir su rol como *manipulanda* en tanto que son las destilaciones de palabras oídas,

habladas, recordadas: objetos concretos, re-identificables y, por lo tanto, manipulables en un sentido básico, común a las demás herramientas.

Otro punto que, en mi opinión, abona una interpretación gradualista de la distinción bajo estudio es la idea de que una de las fuentes importantes de nuestra distancia respecto de otros animales en lo que refiere a capacidad cognitiva es la delegación de la carga cognitiva en el medio ambiente, a través del lenguaje. Este es otro aspecto que conlleva entender las palabras, y el lenguaje en general, como herramientas.

Ahora bien, lo central para nuestra interpretación gradualista es que esta delegación de la carga cognitiva es un recurso que compartimos con muchos otros animales que, como nosotros, dejan marcas distintivas en el ambiente que facilitan su interacción con él, volviéndola más fácil, más rápida o más efectiva. La idea básica es que, para realizar ciertas tareas, algunos organismos realizan marcas reconocibles en el medio ambiental, que así almacena información útil reduciendo el control que de otro modo tendría que ser ejercido por el aparato perceptivo o la memoria del organismo. Así entendida, este tipo de estrategia es fundamentalmente económica:

Una marca de control simplifica el mundo al proporcionarnos una sencilla tarea perceptiva para sustituir a otra de reconocimiento y memorización mucho más difícil (y quizá imposible)<sup>75</sup>.

Aquí vuelve a entrar en juego la interpretación del lenguaje como modalidad perceptiva y de las palabras y conceptos como objetos manipulables que vuelven salientes determinados rasgos útiles del mundo. En esta tarea de economía cognitiva el lenguaje representa no sólo un modo de volver salientes rasgos útiles del mundo sino un modo de hacerlo que a la vez es ordenado, extremadamente sofisticado y de una utilidad, por decir poco, ubicua: esto puede ser así en tanto que los rasgos salientes son a su vez manipulables, son convertidos en nuevos objetos del mundo. Por estos motivos, entre otros, el uso del lenguaje ha disparado infinitamente la capacidad básica de almacenar información en el mundo, cuyo mecanismo en este caso es, sin embargo, esencialmente el mismo que el de las marcas distintivas de los organismos no lingüísticos.

El carácter objetual de las palabras, que las pone en el mismo nivel que el conjunto de herramientas externas, no internalizables, propias de las criaturas

---

<sup>75</sup> Dennett (1996a, p. 162).

gregorianas, no debe desviar nuestra atención de las características centrales que las vuelven herramientas distintivas y fundamentalmente responsables del abismal salto cognitivo que separa al ser humano del resto de los animales. Voy a defender aquí que la acusación de anti-gradualismo hecha al planteo dennettiano puede también responderse apelando a su *explicación* de este salto cognitivo: aunque la diferencia en capacidad cognitiva entre nosotros y los demás sistemas intencionales es innegable, puede no obstante proporcionarse un relato que dé cuenta de ella sin por esto abandonar el modelo explicativo gradualista. Creo que justamente la importancia de esta posibilidad explicativa y el modo cómo algunos puntos de la teoría dennettiana pueden hacerse confluir hacia ella son de algún modo pasados por alto por Clark en su crítica de anti-gradualismo.

En “Belief, Opinion and Consciousness”, Clark reconstruye la propuesta dennettiana sobre contenido mental, tomando como uno de sus ejes la distinción creencia / opinión. El ataque de Clark hacia el rol que ocupa la distinción está dirigido en última instancia contra el anti-realismo dennettiano (o mejor, el aspecto anti-realista de su “realismo moderado”), aunque aquí no me voy a concentrar en esta disputa. Básicamente, el error de Dennett consiste, según Clark, en poner demasiada carga explicativa sobre el rol del lenguaje para dar cuenta de la diferencia cognitiva entre nosotros, seres humanos, y los animales no-lingüísticos, dejando así una brecha inexplicada entre la posesión de creencias y la posesión de opiniones.

Por otra parte, continuaría Clark, una explicación que entienda nuestra capacidad de manipular un lenguaje como producto de una capacidad representacional más básica debida a diferencias de arquitectura computacional (de acuerdo al modelo de “redescripción representacional” de Clark y Karmiloff-Smith<sup>76</sup>) salvaría dicha brecha y constituiría, además, una postura más realista en tanto permitiría utilizar esta diferencia estructural como criterio para los “verdaderos creyentes”; este resultado, a la vez, estaría en consonancia con la intuición básica de sentido común por la cual consideramos que, por ejemplo, un chimpancé tiene creencias y una rana no. El punto de Clark entonces es que el instrumentalismo radical<sup>77</sup> de Dennett, visto desde la perspectiva del par creencia / opinión, no sólo entra en contradicción con nuestras atribuciones de sentido común

---

<sup>76</sup> Clark y Karmiloff-Smith (1993a). La idea central de este modelo es que un “verdadero creyente” (término de Dennett, 1981a) es aquel que posee una organización interna que permite un proceso gradual de redescripción de su propia actividad computacional; este proceso permite al organismo aprovechar la información que ya posea en mecanismos internos de propósito específico, de modo más flexible y accesible a otros dominios cognitivos.

<sup>77</sup> Clark (1990) lo llama “adscriptivismo puro” (*pure adscriptivism*).

respecto de las capacidades cognitivas de diversos organismos, sino que a la vez no proporciona una visión gradualista de dichas capacidades, creando una laguna explicativa inaceptable.

Ahora, el punto general que el ataque clarkiano no toma en consideración es la posibilidad de que una serie de características propias de todos los lenguajes naturales humanos sean, en una medida importante, responsables del desarrollo y de la gradual amplificación de los poderes redescritivos de sus usuarios. La característica que me parece crucial en este sentido es el carácter interpersonal del lenguaje; más específicamente, la idea de que, cada vez que utilizo algún recurso lingüístico particular, me estoy literalmente beneficiando del poder cognitivo que otros han depositado en él en el momento de su fabricación y en el largo proceso de su refinamiento. La posibilidad de beneficiarse de este modo del esfuerzo mental de otros se debe al carácter fuertemente comunicativo que posee el lenguaje.

Estas consideraciones pueden parecer obvias, pero su importancia es central para dar cuenta del modo cómo nuestras mentes humanas se distancian de las demás mentes animales. Dicho de otra forma, una mente individual humana, considerada aislada de un entorno lingüístico, se vería gravemente minusválida, viéndose forzada a coleccionar por sí sola todos los recursos cognitivos para realizar sus tareas. Una mente humana aislada de este modo es de hecho algo que nunca podríamos ver y, por lo tanto, toda vez que tomemos bajo consideración la mente humana para cualquier problema relacionado que estemos tratando, tendremos que tomarla no como un aparato cognitivo cerrado e independiente sino como el resultado de la concentración de grandes cantidades de capacidades, trucos y herramientas acumuladas por el esfuerzo supra-individual de millones de otras mentes similares<sup>78</sup>. Muchas de las capacidades de las mentes humanas están en este sentido unidas a través del medio lingüístico.

Como mencioné antes, en general, mientras más información suponga la fabricación de una herramienta, más inteligencia potencial otorgará a su usuario. Este principio se amplifica exponencialmente en el caso de las “herramientas mentales”, mediadas por el lenguaje, dada la aptitud propia de las palabras para su fácil transmisión, aptitud relacionada nuevamente con su carácter objetual que, como vimos, liga las raíces del pensamiento a una modalidad perceptiva. Las palabras, en este sentido, son un tipo de herramienta que incorpora información acumulada a través de

---

<sup>78</sup> Se trata, más aún, de una extensión sincrónica –con los contemporáneos– y diacrónica –con el pasado cultural.

una transmisión cultural de miles de años entre miles de millones de mentes: este altísimo nivel de sofisticación y perfeccionamiento implica una acumulación de inteligencia potencial sin igual, comparada con cualquier otro tipo de herramienta diseñada por animales no lingüísticos.

Ahora bien, no sólo debemos pensar en el beneficio interpersonal que otorga el lenguaje teniendo en cuenta las palabras como herramientas individuales, sino que el beneficio que aporta el lenguaje debe entenderse también, en general, en términos de trucos y estrategias para el mejoramiento del rastreo, la colección y el aprovechamiento (posiblemente novedoso) de la información, esto es, en sentido amplio, como un “aprender cómo pensar mejor”. En último término, ésta es una característica general que el lenguaje comparte con otras herramientas con beneficio cognitivo:

Las criaturas gregorianas dan un gran paso hacia el nivel humano de destreza mental, beneficiándose de la experiencia de otros al explotar la sabiduría encarnada en las herramientas mentales que esos otros han inventado, mejorado y transmitido; por eso aprenden cómo pensar mejor sobre aquello que tienen que pensar a continuación... y así sucesivamente, creando una torre de sucesivas reflexiones internas sin límite fijado ni discernible<sup>79</sup>.

En este sentido, el aprendizaje del lenguaje implica en el fondo la incorporación de toda una estructura pre-diseñada que reúne no sólo la información acumulada en las herramientas individuales que la componen sino también la información implícita en el sistema de relaciones entre esas herramientas, lo cual dispara enormemente nuestra habilidad redescritiva innata. La crítica de Clark pierde fuerza si en alguna medida se acepta la inversión del orden de prioridad aquí explorada: es decir, el enorme poder redescritivo que las estructuras lingüísticas otorgan por sobre nuestras capacidades innatas<sup>80</sup>.

Hay otros problemas en la crítica clarkiana. He mantenido que una diferencia central entre las palabras y otras herramientas es el poder que aquellas tienen para acumular información heredada a lo largo del proceso de su conformación y depuración. Las palabras tienen este poder en tanto que, como ya destacamos en diversos puntos, son elementos relativamente fijos, con un contenido relativamente estable. En este sentido, el contenido de las opiniones, en tanto son estados que resultan de un proceso

---

<sup>79</sup> Dennett (1996a, p. 123).

<sup>80</sup> La presente discusión tampoco debiera interpretarse estrictamente en términos de la disputa innato / adquirido (disputa que puede llegar a ser menos importante, y más engañosa, de lo que es considerada) sino más bien como una disquisición sobre la fuente del poder cognitivo que el lenguaje otorga.

de escrutinio por medios lingüísticos, puede considerarse infinitamente más determinado y preciso comparado con el de las meras creencias. Las opiniones, de hecho, incorporan directamente el carácter fijo y determinable de las palabras que las componen.

Por otra parte, la creencia, tal como es entendida en el contexto de la teoría de los sistemas intencionales, es esencialmente un estado mental abstracto, con lo cual el problema de la determinación de su contenido es un problema derivado, propio del intérprete, y cuya resolución nunca es concluyente en tanto que depende en una importante medida de los fines explicativos y predictivos del intérprete. Las opiniones, para poder conformarse como tales, deben conllevar algún tipo de expresión lingüística, sea ésta de hecho proferida verbalmente, sea mantenida tácita en el ensayo mental que típicamente identificamos con el fenómeno del pensamiento conciente; por esto, uno supondría que siempre se puede (al menos en principio) apelar al agente interpretado para confirmar exactamente el contenido de su estado de opinión.

Dado este planteo, lo que quiero sostener es que la interpretación de Clark caracteriza erróneamente la diferencia entre las creencias y las opiniones, confundiéndola con la diferencia trivial que habría, y que tanto Dennett como Clark admiten, entre los criterios para la identificación y determinación de los contenidos mentales de las opiniones y los de las creencias. Un problema general en la interpretación que Clark hace de la postura dennettiana es asumir que la motivación primordial de su anti-realismo es la defensa de la tesis de la indeterminación del contenido de los estados mentales:

El “instrumentalismo” de Dennett o su negación a abrazar un realismo pleno puede en buena medida rastrearse en su rechazo a aceptar una visión en la que habría hechos determinados respecto de qué adscripción de actitud proposicional captura precisamente nuestro estado mental<sup>81</sup>.

La importancia de la tesis de la indeterminación del contenido mental para el realismo moderado de Dennett no es, de hecho, menor; pero tampoco es central. Mi postura (ver pp. 9, 15-16) ha sido que la fuente teórica de este tipo de anti-realismo está implícita en la naturaleza del mismo mecanismo de la actitud intencional, tal como es entendido por su autor: un mecanismo cuyos productos son necesariamente idealizados,

---

<sup>81</sup> Clark (1990, p. 144). La traducción es mía.

aproximados, difusos, aunque útiles, fundamentalmente rasgos no susceptibles de dar lugar a una interpretación estrictamente realista.

La mala reconstrucción del anti-realismo dennettiano influye, de modo derivado, en la interpretación clarkiana de la distinción creencia / opinión, tomada como eje del planteo general de Dennett. En este sentido, Clark reconstruye la distinción como un modo de ordenar la respuesta a un tipo particular de realismo, el realismo atomista fodoriano, que muy marcadamente se deja engañar por el carácter determinado de nuestras adscripciones mentalistas de sentido común y las toma como los correlatos de estados internos con contenido determinado<sup>82</sup>. Lo que sostengo es que la distinción entre creencias y opiniones no es meramente un recurso para poner de manifiesto la ilusión de determinación de los contenidos mentales. Éste es un papel que sin lugar a duda la distinción cumple y en un campo temático al que Dennett ha dedicado un esfuerzo argumentativo importante. De todos modos, creo que dicho papel no es único ni central para la distinción.

Mi interpretación es que la distinción funciona principalmente (pero tampoco exclusivamente) en el marco del programa dennettiano de definir y describir la conformación de lo que Dennett (1996a) llama diferentes “tipos de mentes”, entendidas como conglomerados abstractos de disposiciones y atribuciones intencionales<sup>83</sup>; en este sentido, y aquí es donde el concepto de opinión entra en juego, quiere en especial dar cuenta del modo cómo el lenguaje tiene un efecto retrospectivo que altera profundamente nuestra particular mentalidad humana. Clark no logra interpretar la importancia y consecuencias de la determinación adquirida con las opiniones<sup>84</sup>, esto es, el carácter fijo señalado, determinación que justamente habilita su rápida manipulación e intercambio, lo que a su vez genera una compleja estructura de conceptos e ideas, que gradualmente va configurando el conjunto de hábitos mentales constitutivos de la mente conciente de un ser humano adulto lingüísticamente competente.

Por otra parte, es importante también señalar que Dennett, justamente por su inclinación gradualista, nunca negaría (ni mucho menos) la existencia de complejidades estructurales innatas mayores en el caso humano que en el de los demás animales. Esto es, para la defensa de las tesis anteriores no se precisa negar la necesidad de un

---

<sup>82</sup> Este punto será tratado en detalle en el capítulo 4.

<sup>83</sup> Esto tiene su contraparte meta-teórica, a la que aludo en p. 124 (nota 92).

<sup>84</sup> Cabe agregar que, en “Belief, Opinion and Consciousness”, Clark de hecho caracteriza repetidamente las opiniones como estados “relativamente superficiales” y se refiere a su contenido como “trivialmente determinado”.

complejo aparato interno que pueda dar cuenta del aprendizaje de cualquier lenguaje natural. Más aún, tampoco sería estrictamente necesario negar lo que Clark y Karmiloff-Smith afirman, esto es, la existencia de predisposiciones innatas para el hábito de la redescipción representacional. El punto aquí es evaluar la medida en que el surgimiento paulatino del lenguaje, entendido como un conjunto de herramientas, mañas y trucos acumulados, junto con la facilidad de su transmisión, pueda haber influido en la conformación de esas predisposiciones.

La posibilidad evolutiva de esta transferencia de habilidades al genoma humano se denomina “efecto Baldwin”, proceso por el cual los descubrimientos a nivel fenotípico de nuevas habilidades crean a la vez nuevas presiones selectivas que tenderán a favorecer los organismos con configuraciones estructurales más cercanas a las que posibilitaron dichos descubrimientos<sup>85</sup>. No me extenderé aquí en este punto, pero habría que tener en cuenta su posibilidad teórica a la hora de buscar un asiento biológico para las ideas dennettianas aquí tratadas.

### ***3. Relevancia e Implicaciones de la Distinción***

Espero haber mostrado que el carácter gradualista de la distinción entre creencias y opiniones descansa mayormente, y de modo paradójico, en la explicación de la diferencia radical que separa los dos términos y que hace de las opiniones las responsables del inmenso salto cognitivo entre el hombre y los demás organismos vivos. Es, por lo tanto, de modo acorde a la tendencia gradualista (y evolucionista) general de la teoría de los sistemas intencionales que se produce un quiebre teórico fuerte entre dos tareas de investigación diferentes, quiebre que responde a la distinción fuerte entre creencias y opiniones. El quiebre es el siguiente:

Mi pronóstico empírico es que una psicología cognitiva apropiada va a tener que trazar una distinción muy tajante entre creencias y opiniones, que la psicología de las opiniones realmente va a tener que ser muy diferente de la psicología de las creencias, y que los tipos de arquitectura que funcionarán muy bien para, digamos, las creencias perceptuales no lingüísticas [...] van a tener que ser sustituidos sustancialmente para poder manejar las opiniones<sup>86</sup>.

---

<sup>85</sup> Dennett explora la idea del efecto Baldwin en *Consciousness Explained* (1991a, pp. 197 y siguientes).

<sup>86</sup> Dennett (1998a, pp. 89-90). La traducción es mía.

La distinción en cuestión se ubica, así, en el seno de una decisión metodológica fundamental para el planteo dennettiano. A la vez, dado, por una parte, el gradualismo de la distinción y dada, por otra parte, esta importante ramificación metodológica, puede explorarse la medida en que la propuesta de una estrategia disyuntiva para los problemas relacionados al contenido mental y la intencionalidad tiene repercusiones gradualistas a nivel metodológico, en tanto que la incorporación de la distinción en las discusiones pertinentes puede poner a la luz planteos, argumentos, rodeos y confusiones últimamente anti-gradualistas o anti-evolucionistas. Habiendo ya explorado las razones que vienen en contra de una interpretación anti-gradualista de la distinción en tanto planteo sustantivo (de nuevo, relacionadas mayormente con el modo cómo la incorporación del lenguaje puede explicar el salto cognitivo humano), quiero ahora definir los roles metodológicos que ésta cumple, dado su gradualismo, en el contexto de la teoría de los sistemas intencionales y frente a otras teorías del contenido que no incorporan la propuesta dennettiana.

Quiero sostener que la distinción bajo estudio es un punto básico para entender el planteo general dennettiano y, por este motivo, para dar cuenta del modo cómo este planteo se enfrenta a otros planteos rivales. Esto es, el “pronóstico empírico” antes citado no es algo circunstancial o menor, sino que responde directamente a aspectos centrales de la postura dennettiana sobre contenido mental. En este sentido, las raíces de una defensa de la estrategia propuesta pueden encontrarse en algunos aspectos centrales de la teoría de los sistemas intencionales, especialmente el modo en que ésta circunscribe los estados intencionales como objeto teórico.

Como vimos en el apartado 1.1, la creencia, el deseo y los demás estados intencionales son discernibles, a partir de la adopción de la actitud intencional, tomando en consideración lo que llamaba la “situación ambiental” del organismo, es decir, el conjunto hipotético de todas las interacciones, presentes y pasadas, de un organismo con su medio ambiental. La interacción de un organismo con el medio ambiental en que se encuentra puede redesccribirse en términos del *flujo de información* que el organismo recibe del medio y de sus reacciones ante esa información. La distinción creencia / opinión puede, a su vez, incorporarse a este modo de enfocar la “situación ambiental” del organismo en la puesta en práctica de la estrategia intencional. En lo que sigue me voy concentrar en este enfoque particular.

El flujo de información relevante para la atribución del estado de creencia (u otros estados intencionales) es, fundamentalmente, la información perceptual

proveniente del ambiente a la que, podemos suponer, responde el organismo. El organismo interpretado reacciona frente a los estímulos sensoriales del medio y nuestra atribución intencional relaciona esas reacciones con el rol y los objetivos estipulados para aquel. Para el caso de las opiniones, el flujo de información sigue siendo la mencionada base perceptual, pero con el agregado obligado de la información encapsulada en las ideas y conceptos lingüísticamente arraigados<sup>87</sup> que se estén considerando.

No es necesario repetir que el tipo de información y los mecanismos de su procesamiento que aquí entran en juego son fundamentalmente distintos de los primeros, tanto que puede pensarse en dos capas distinguibles de procesamiento. Este solapamiento entre dos flujos distintos de información tiene consecuencias importantes para una teoría del contenido mental. Hay dos en particular que quiero destacar: la primera es una visión de la creencia como sustrato persistente, condicionado por la información perceptual, y la segunda es la relación directa que la creencia mantiene con la conducta<sup>88</sup>.

Respecto de esto último, quiero agregar que, dado el papel fundamental que juega la observación de la conducta en el mecanismo de la estrategia intencional, el carácter más o menos directo de la relación que podamos trazar entre dicha observación y nuestra atribución intencional es de suma importancia, incluso para la delimitación del objeto de estudio para el cual la estrategia intencional fue postulada. Las creencias, por definición, son determinadas *directamente* apelando a la conducta del organismo interpretado. Por otra parte, las opiniones no son confiables para la predicción de la conducta en tanto que el modo en que adquirimos, mantenemos y abandonamos opiniones es generalmente laxo y, por esto mismo, más o menos independiente de nuestras predisposiciones conductuales.

En este sentido, la manera en que intercambiamos opiniones es absoluta e imperceptiblemente “*liberal*”: no regido por pautas fijas, impreciso, desordenado, muchas veces no del todo conciente y por momentos incoherente. Refiriéndose a esta liberalidad del comercio de opiniones, Dennett dice:

---

<sup>87</sup> Hago alusión aquí a la idea de las palabras como predecesoras de los conceptos, considerada en pp. 53-54.

<sup>88</sup> Mencioné este segundo punto en p. 50.

Podemos heredarlas, entrar en posesión de ellas sin darnos cuenta, no descartarlas después de haber decidido hacerlo, tomarlas en préstamo temporal y olvidar que esto es lo que hemos hecho<sup>89</sup>.

Y:

A veces conservamos una oración porque nos gusta el sonido que tiene, o porque seremos luego recompensados por reproducirla ante una solicitud, o sólo porque tiene una suerte de poder de persistencia [*staying power*] en nuestra imaginación<sup>90</sup>.

La relación indirecta entre conducta y opiniones se vincula con la segunda característica que destacaba. La creencia puede considerarse un *sustrato persistente* respecto de las opiniones en tanto que en ella únicamente influye la información perceptual del ambiente. La idea de las palabras como constituyendo una nueva modalidad perceptiva quiere marcar una escisión entre dos maneras diferentes en que incorporamos información, una de las cuales es más básica que otra y, a la vez, es suficiente para que el organismo sea susceptible de interpretarse desde la actitud intencional<sup>91</sup>. Toda la información que intervenga por encima del estrato más básico de información sensorial puede considerarse adicional y, en este sentido, secundaria para el estudio de la creencia y los demás estados intencionales.

Lo que quiero destacar en base a las dos características mencionadas es que la circunscripción al fenómeno de la creencia (término técnico) es justificable como una *toma de posición metodológica* para una teoría del contenido, en el siguiente sentido: la creencia puede entenderse como una instancia básica hacia la cual pueden hacerse confluir las preguntas y los problemas básicos de una teoría del contenido, mientras que la opinión puede entenderse como un caso derivado, especial de creencia, que incorpora problemas adicionales relacionados específicamente con el papel del lenguaje en la conformación de las opiniones. En otros términos, las opiniones pueden concebirse como parte de un sistema “despegado” del de las creencias, con problemas propios e independientes.

Algunas afirmaciones de Dennett parecen confirmar mis ideas sobre esta asimetría en el núcleo del planteo metodológico dennettiano:

---

<sup>89</sup> Dennett (1978a, p. 307). La traducción es mía.

<sup>90</sup> Dennett (1978a, p. 47). La traducción es mía.

<sup>91</sup> En realidad, como se explicó en el apartado 1.1, no hay una restricción o requisito en el agente para la aplicabilidad de la actitud intencional, aunque su uso justificable implica, en la práctica, cierta variabilidad conductual que, como mínimo, debe responder a su vez a alguna forma de captación de estímulos perceptuales. Hay una mención de este punto en Dennett (1969, p. 106).

Al ser criaturas parlantes, es inevitable que, a menudo, lleguemos a creer que una oración determinada, realmente formulada, deletreada y puntuada es verdadera, y que en otras ocasiones lleguemos a creer que esa oración se haga realidad, pero éstos son *casos especiales* de creencia y deseo y como tales pueden no ser *modelos fiables* para la totalidad del campo de acción<sup>92</sup>.

En otra parte, Dennett afirma que las opiniones “no son estados *ejemplares* o *fundacionales* de creencia” y, continúa, “[s]i uno comienza, como debería, con los estados y eventos cognitivos ocurientes en animales no humanos, y los usa como el *fundamento* sobre el cual construir teorías de la cognición humana, se verá más fácilmente que los estados lingüísticamente infectados [esto es, las opiniones] son derivados [y] menos directamente implicados en la explicación de la conducta...”<sup>93</sup>.

Creo que esta separación tajante al nivel del planteo de la investigación puede extenderse a otra separación temática más general que estructura la obra dennettiana en su conjunto: esto es, la separación entre una teoría del contenido mental y una teoría de la conciencia<sup>94</sup>. La primera abordará problemas generales sobre la intencionalidad y la atribución de contenidos mentales desde la perspectiva básica de la estrategia intencional, mientras que la segunda abordará específicamente el funcionamiento y las derivaciones de la práctica culturalmente embebida de la atribución de opiniones, estados cuya contribución es primordial para la constitución de la conciencia, en algunas de sus acepciones centrales<sup>95</sup>.

Si la circunscripción de la teoría del contenido mental al fenómeno de la creencia es justificada, entonces puede mantenerse que la distinción propuesta entre creencias y opiniones responde, entre otras cosas, a una fuerte motivación meta-filosófica, esto es, la demarcación de áreas temáticas diferentes en un campo de estudio altamente heterogéneo. Esto implica, en definitiva, que la defensa de la distinción creencia / opinión excede su rol teórico sustantivo.

Espero haber mostrado hasta aquí que la importancia que tiene la distinción se ramifica hacia los múltiples niveles en los que funciona: la distinción es, en primer término, parte de un esfuerzo que apunta en general hacia la “de-lingüistificación” de lo

---

<sup>92</sup> Dennett (1987a, p. 32). El subrayado es mío.

<sup>93</sup> Dennett (1998a, p. 363). La traducción y el subrayado son míos.

<sup>94</sup> Esta es una de las perspectivas más arraigadas en la teoría general de Dennett, reflejada ya en el título de su primer libro, *Content and Consciousness (Contenido y Conciencia)*, de 1969.

<sup>95</sup> Dennett no es muy cuidadoso en diferenciar lo que en la literatura han sido reconocidos como diferentes tipos de conciencia, algunos de los cuales podrían quizás ser aplicables a animales no lingüísticos, pero su utilización central del término refiere exclusivamente al tipo de conciencia atribuible de modo paradigmático a seres humanos lingüísticamente competentes.

intencional y de lo mental, esto es, hacia un tratamiento filosófico de dichos temas que no esté enfocado a través del lente que proporcionan las descripciones mentalistas cotidianas, mediadas por el lenguaje y la cultura<sup>96</sup>; en segundo término, como veíamos en el anterior apartado, la distinción constituye el eje estructural para el estudio de la conformación de diferentes tipos de mentes, teniendo especialmente en mira el caso humano, en el que se apunta a dar cuenta de la transformación que, *via* lenguaje, sufren las creencias y los demás estados mentales; por último, la distinción funciona meta-filosóficamente como “separadora de aguas” entre temáticas cuya relación, según el planteo dennettiano, es fuertemente indirecta, tales como las referidas al contenido mental y la intencionalidad, por un lado, y las referidas al problema de la conciencia, por otro. Dado el carácter asimétrico de la distinción destacado antes, su importancia meta-teórica reside justamente en entender los problemas de la conciencia como problemas fuertemente *derivados* y que, justamente por este motivo, merecen un tratamiento aparte al de los problemas de la intencionalidad, con categorías y herramientas conceptuales diferentes.

Enfocar la distinción creencia / opinión desde la perspectiva del modo cómo Dennett entiende la psicología popular o de sentido común nos proporcionará una visión renovada de su rol y de sus implicaciones metodológicas, que será, además, la clave de la discusión entre las teorías dennettiana y fodoriana en el próximo capítulo. En particular, tengo en mente las ideas que Dennett propone en el artículo “Two Contrasts: Folk Craft versus Folk Science, and Belief versus Opinion” (1991d) y que vienen a complementar la posición general sobre la psicología popular defendida en *The Intentional Stance*, además de profundizar sobre la distinción aquí estudiada.

En el mencionado artículo, Dennett propone entender la psicología popular no tanto como una teoría sino como una habilidad (*craft*), pero a la vez reconociendo la existencia de un fenómeno paralelo a dicha habilidad aunque fundamentalmente diferente de ella: lo que llama la “ideología” de la habilidad o, de acuerdo con el título del artículo, la “ciencia popular” que versa sobre ella. La psicología de sentido común debería entenderse, según Dennett, más bien como un talento puesto en juego en la práctica, algunos de cuyos elementos podrían ser innatos. Este talento debe diferenciarse claramente del discurso heredado que refiere a él y que comporta la postulación de diversas entidades, procesos y propiedades.

---

<sup>96</sup> Este es el rol de la distinción destacado principalmente tanto por Clark como por Frankish.

Dado esto, propongo trazar una analogía entre esta distinción y la distinción creencia / opinión. Del mismo modo en que, teniendo en cuenta la influencia del lenguaje, pueden distinguirse dos modalidades de fenómenos intencionales, una básica y otra derivada, también, y de acuerdo con el mismo criterio, pueden distinguirse dos niveles sociales en los que dichos fenómenos son atribuidos. A la vez, tal como había sido el caso para el par creencia / opinión (pp. 46-47), la separación tajante entre habilidad e ideología, en el caso de seres eminentemente lingüísticos como nosotros, es en alguna medida una herramienta o un artificio teórico en tanto que la habilidad en cuestión, como es natural, se proyecta a través del comportamiento lingüístico.

Por otro lado, el gradualismo implícito en la circunscripción de las opiniones al caso humano y en el correspondiente enfoque de la teoría dennettiana en los sistemas intencionales no lingüísticos (estudiado en el apartado 3.2) tiene su contraparte en la psicología popular. La modalidad más básica de esta última, esto es, en tanto que es entendida como habilidad, es la que en principio puede atribuirse a organismos no lingüísticos y, de hecho, está siendo estudiada y confirmada entre los grandes simios<sup>97</sup>. Esta es una manera de entender la psicología popular como una actividad natural, no dependiente de una teoría previa y cuyos orígenes pueden rastrearse evolutivamente en algunas de las características que los seres humanos compartimos con otros primates, especialmente las relacionadas con su sociabilidad.

Por otra parte, mientras pueda entenderse que la psicología popular en este sentido opera, en una primera instancia, sobre las creencias en base a la conducta no lingüística, la psicología popular atravesada por su ideología culturalmente heredada operaría sobre las opiniones estrictamente mediante el reporte lingüístico. Desde esta perspectiva, y teniendo en cuenta el aducido carácter derivado de las opiniones y la “liberalidad” de su intercambio (pp. 62-63), el discurso mentalista cotidiano cobra un matiz particular, que, como veremos, resultará ser de inmensa importancia para el teórico del contenido mental:

Postulamos todas estas actividades y procesos mentales aparentes para encontrarle sentido a la conducta que observamos, para, en realidad, encontrarle el mayor sentido posible a la conducta, especialmente cuando la conducta que observamos es la nuestra propia. [...] [C]ada uno de nosotros es, en muchos aspectos, una especie de auto-

---

<sup>97</sup> Hay una consolidada línea de investigación en etología cognitiva (ver Whiten (1996) para un enfoque sinóptico) que tiene su origen en el artículo fundacional de Premack y Woodruff de 1978, “Does the Chimpanzee Have a Theory of Mind?”. En esta línea se utiliza el término “teoría de la mente” de modo pre-crítico, sin implicaciones para el contraste propuesto por Dennett.

psicólogo inveterado, que inventa sin esfuerzo interpretaciones intencionales de nuestras propias acciones en una mezcla inseparable de confabulación, autojustificación retrospectiva, y (a veces, sin duda) buena teorización<sup>98</sup>.

El núcleo de la actitud intencional, esto es, la asunción optimizadora de racionalidad, se ramifica para el caso de las opiniones en la producción de relatos complejos que incorporan los conceptos de la ideología (las entidades, procesos y propiedades postuladas), que a su vez fueron forjados en el contexto de esta actividad racionalizadora.

Es desde esta perspectiva que podemos entender más claramente el sentido en que, para Dennett, como veremos en el siguiente capítulo, las categorías de la psicología popular en tanto que ideología pueden ser engañosas en cuanto se las pretenda tomar como referentes para la investigación científica sobre la mente. En acuerdo con esto, podremos interpretar el análisis del lenguaje mentalista con el que Dennett se arma (y del cual el concepto de “opinión” es una herramienta central) como un prerrequisito de su teorización y, a la vez, como una fuente de críticas y contra-argumentos contra las teorías que, al asumir los supuestos y las distinciones que tal lenguaje incorpora, se modelan sobre él.

Este es (aproximadamente) el rol para la distinción creencia / opinión que resalta más claramente de los textos dennettianos, y que es generalmente reconocido en la literatura; a pesar de esto, espero haber mostrado a lo largo de este capítulo que dicho rol se enraíza en aspectos absolutamente básicos del planteo de Dennett. El éxito en esta tarea es un paso importante hacia el objetivo principal de este trabajo: trazar el vínculo profundo que la teoría de los sistemas intencionales mantiene entre, por un lado, la defensa de la distinción entre creencias y opiniones y, con ella, de la necesidad de un análisis del lenguaje mentalista de sentido común (como contrario a la asunción pre-crítica de sus categorías) y, por otro lado, su carácter fuertemente evolucionista y gradualista.

Quiero, finalmente, poner en claro un punto importante respecto de mi interpretación de los términos de la distinción. Evidentemente, la aceptación del múltiple rol teórico para la distinción que aquí he defendido ya presupone su interpretación extendida, discutida en el apartado 3.1. Pero hay una aclaración adicional que hacer: estoy tomando la distinción en un sentido amplio que distingue en general

---

<sup>98</sup> Dennett (1987a, p. 89).

los distintos procesos cognitivos en los que interviene el lenguaje de aquellos no lingüísticos.

De este modo, por una parte, interpreto el término técnico “creencia” como exponente de los estados intencionales disposicionales en general (creencias, deseos, intenciones, esperanzas, dudas, etcétera) y, por otra parte, interpreto el término técnico “opinión” como exponente de todo proceso interno en el que influya de algún modo el lenguaje, incluidos el conjunto de fenómenos que asociamos algo difusamente con el problema de la conciencia (como, por ejemplo, la experiencia sensorial, la introspección y el conocimiento en primera persona). Aunque en la escasa presentación dennettiana de la distinción no hay estrictamente puntos de apoyo para esta extensión, una interpretación más abarcadora es justificada si atendemos al vínculo entre el planteo metodológico general de la teoría de los sistemas intencionales y las acusaciones que Dennett, apelando a la distinción, hace a otras teorías rivales.

Dichas acusaciones, de hecho, aunque generalmente formuladas en el contexto de argumentaciones sobre problemas puntuales, son en el fondo de muy largo alcance: considerando lo que aquí he estado defendiendo, es decir, la importancia que la distinción tendría, en general, para el planteo de la teoría de los sistemas intencionales y, en particular, para la delimitación justificada de su objeto de estudio, tenemos una situación en la que la incorporación en una teoría del contenido mental de la distinción entre creencias y opiniones, o alguna distinción afín pero con similares consecuencias teóricas y metodológicas, se convierte en un punto central para medir la compatibilidad o incompatibilidad de cualquier teoría frente al planteo dennettiano. Dicha compatibilidad o incompatibilidad podrá rastrearse desde el nivel de las discusiones y los argumentos particulares hasta el de los planteos teóricos generales, lo que incluye el conjunto de presupuestos en los que aquellos se apoyan. De esta manera, el presente análisis del rol amplio que cumple la distinción podrá considerarse de aquí en adelante como una *hipótesis a ser confirmada*. En el próximo capítulo me dedicaré a dicha confirmación, que se revelará del mismo modo tanto al nivel de las discusiones particulares como al de los presupuestos teóricos y metodológicos subyacentes.

---

## El Alcance del Enfoque Evolucionista

### *1. Concepciones Pragmatista e Intelectualista de la Psicología Popular*

Como mencionaba al concluir el anterior capítulo, creo que las acusaciones que Dennett hace a algunas teorías rivales apelando a la distinción entre creencias y opiniones, tienen un alcance muy profundo; un alcance que, a mi parecer, puede rastrearse en última instancia hasta el carácter metodológico general que en los capítulos 1 y 2 he denominado “enfoque evolucionista acerca del contenido mental”, en el caso de la teoría de los sistemas intencionales, y “enfoque anti-evolucionista acerca del contenido mental”, en el caso de la teoría representacional de la mente. Esta idea será desarrollada y fundamentada a lo largo del presente capítulo.

La perspectiva argumentativa que adoptaré será la de relevar las implicaciones teóricas y metodológicas que se siguen, en el contexto de la disputa entre las propuestas de Dennett y Fodor, de incorporar (o no) una distinción entre creencias y opiniones, tal como la he presentado y especialmente considerando su gradualismo, y de establecer el modo cómo dichas implicaciones se integran a los enfoques evolucionista y anti-evolucionista de dichas propuestas. Para esto, me concentraré en los que, creo, son los puntos encontrados más relevantes entre ambos autores: la interpretación del discurso psicológico de sentido común, la pertinencia de un análisis filosófico basado en el concepto de “actitud proposicional”, la cuestión del realismo intencional y la relación entre el nivel personal y el sub-personal de explicación. Todas estas discusiones están, como se verá, íntimamente relacionadas, y el punto común de enfoque será el rol que en ellas juega, o que podría jugar si se incorporara, la mencionada distinción. En el presente apartado, me concentraré en la psicología popular.

Tanto la teoría de los sistemas intencionales como la teoría representacional de la mente se asientan, de un modo u otro, en un análisis del discurso mentalista de la psicología de sentido común. Tanto Dennett como Fodor aceptan esto explícitamente<sup>99</sup>. Esto es, el punto de partida que ambos autores aceptan y consideran justificadamente

---

<sup>99</sup> Ver Dennett (1987a, p. 52) y Fodor (1987, p. 28).

pertinente para la elaboración de sus teorías de la intencionalidad es compartido: el hecho asumido por ambos es que términos intencionales como “creer”, “desear” y afines pertenecen al lenguaje común y que, por lo tanto, deberemos recurrir en una primera instancia a él (y al uso cotidiano de dichos términos) para dar cuenta de ellos y, especialmente, para dar cuenta del papel que tendrían que jugar en una psicología científica.

Este es un hecho muy relevante, que no debiera pasar inadvertido, y que refleja por lo menos tres cosas: (1) ambos autores consideran que el discurso intencional cotidiano es identificable como fenómeno particular y con características propias; (2) dado el punto 1, ambos autores atribuyen alguna relevancia teórica al análisis de dicho discurso y sus características; (3) ambos autores comparten la tesis implícita de que el marco adecuado para la discusión filosófica sobre el carácter intencional de la mente, esto es, el hecho de que pueda “ser acerca de” otra cosa, es dicho discurso (o, por lo menos, que la referencia a él es ineludible para la discusión). Una consecuencia muy importante de todo esto, para tener especialmente en cuenta en lo que sigue, es que (casi) todos los problemas a los que se enfrentarán las teorías dennettiana y fodoriana, así como su justificación, dependerán en mayor o menor medida de la particular concepción de la psicología popular que las teorías asuman o defiendan explícitamente<sup>100</sup>.

Ya esboqué en los capítulos 1 y 2 los rasgos básicos de las concepciones o “reconstrucciones” de la psicología popular que Dennett y Fodor respectivamente defienden. Llamaré a la concepción defendida por el primero “*concepción pragmatista de la psicología popular*” y a la defendida por el segundo “*concepción intelectualista de la psicología popular*”<sup>101</sup>. Aunque haya puntos de acuerdo, las posturas son en general opuestas, y creo que puede darse cuenta de esta oposición apelando a la distinción entre creencias y opiniones (a su incorporación en el primer caso y a su ausencia en el segundo).

Ambas posturas son explícitamente anti-eliminativistas respecto del discurso psicológico popular y, de acuerdo con esto, defienden dos puntos relacionados: (1) hay características propias de dicho discurso en virtud de las cuales éste aporta un poder

---

<sup>100</sup> A la vez, éste es el motivo principal por el que todos los puntos que trataré a lo largo de este capítulo están en íntima conexión.

<sup>101</sup> Cabe aclarar que estas denominaciones no son asumidas por ninguno de los autores o de sus comentaristas y, más aún, que ni Dennett ni Fodor defienden explícitamente una “concepción” acerca de la psicología popular.

predictivo y explicativo irreducible a vocabularios puramente no intencionales, y (2) dicho discurso es prácticamente indispensable dada su contribución a la interpretación de la conducta de las personas y de otros organismos<sup>102</sup>. Ahora bien, siendo esencialmente éstos los puntos de acuerdo, las concepciones en cuestión divergen de modo radical en cuanto a la explicación del poder predictivo de la psicología popular o, en otros términos, a la identificación de sus características propias (punto 1), características que a la vez dan cuenta de su indispensabilidad práctica (punto 2). Me concentraré primero en la concepción pragmatista de la psicología popular.

Hacia el final del capítulo 3 (p. 66) había propuesto extender el análisis dennettiano del discurso mentalista a partir de los conceptos de “habilidad” e “ideología” a la distinción entre creencias y opiniones. Es en este sentido que propongo enfocar la disputa entre Dennett y Fodor respecto de la psicología de sentido común en base a esta última distinción. En primer lugar, quiero señalar que, si se acepta que la concepción pragmatista se estructura de este modo disyuntivo (a partir de los pares habilidad / ideología y creencia / opinión), entonces no puede ver a la psicología popular como un fenómeno homogéneo. Creo que, siendo un producto popular, hay razones *prima facie* (y en contra de Fodor<sup>103</sup>) para entender la psicología de sentido común más bien como una mezcla. Aún así, pueden distinguirse aspectos (o conjuntos de aspectos) en ella, y éstos se corresponderían aproximadamente a los destacados a través de la distinción entre creencias y opiniones.

De acuerdo con esto, la psicología popular consta de un nivel básico definido esencialmente en términos disposicionales: el discurso mentalista *es*, según la concepción pragmatista, fundamentalmente una habilidad (*craft*). Esto debe entenderse dentro del contraste explícito con el concepto de ideología. La idea aquí es que aprendemos a usar el discurso mentalista en las circunstancias apropiadas y con esto incorporamos fundamentalmente la aplicación adecuada de los términos intencionales, pero sin embargo no aprendemos en este proceso a qué refieren dichos términos (no aprendemos explícitamente qué es una creencia, qué es un deseo, qué es una impresión, etcétera): es decir, no aprendemos primeramente la ideología que teoriza sobre la habilidad, ni para adquirir esta última es necesario apelar a la primera. Con palabras de Dennett:

---

<sup>102</sup> Un tercer punto anti-eliminativista refiere al papel del discurso mentalista en una psicología científica y será uno de los ejes de la discusión en el apartado 4.4.

<sup>103</sup> Retomo esta crítica en pp. 76-77.

Desde luego que nadie *nos dice* qué son las creencias, o si alguien lo hace, o por si casualidad especuláramos sobre el tema por cuenta propia, la respuesta a la que llegamos, sabia o tonta, se traslucirá en forma muy débil en nuestros hábitos de pensar sobre lo que la gente cree<sup>104</sup>.

Es importante tener en cuenta el carácter asimétrico de las distinciones que la concepción pragmatista asume. Esto es: por más diversificado y heterogéneo el fenómeno de la psicología popular pueda de hecho ser, muy especialmente para el caso de sus manifestaciones más desarrolladas y sofisticadas, su núcleo y su aspecto más relevante para el análisis filosófico se revela en la práctica como un “saber-cómo” ryleano, un conjunto difuso de disposiciones, una capacidad inespecífica para ser reactivo, de determinadas formas en determinadas situaciones, frente al comportamiento de otros. El conjunto de información adicional, especificada por medios lingüísticos y heredada culturalmente de modo difuso (impreciso y no necesariamente conciente), es considerada, cuanto menos, secundaria.

Ahora bien, para el caso estándar de las atribuciones intencionales entre personas adultas lingüísticamente competentes, y en la medida en que se considere que la expresión lingüística es un medio concreto en que la habilidad se manifiesta y desarrolla, entonces la psicología popular estará a lo sumo *respaldada* por el conjunto culturalmente mediado de opiniones sobre la habilidad. Este conjunto de opiniones puede ser información muy relevante para comprender la concepción popular de la mente o la interpretación del vocabulario intencional más arraigada en un grupo social determinado, pero no es constitutivo de, ni necesario para, la psicología popular como fenómeno global y básico, común a toda persona.

La concepción pragmatista del discurso psicológico cotidiano, según la cual sus características más relevantes responden a necesidades contextuales y se interpretan dentro de un ámbito social concreto, de por sí favorece la visión anti-cientificista de la psicología popular que, por otros argumentos, Dennett defiende de modo explícito. Hay que recordar que (de acuerdo a mi interpretación de la teoría de los sistemas intencionales, en el apartado 1.1) el núcleo de la defensa explícita de esta visión radica en la tesis dennettiana de que existe un aspecto normativo irreductible en el mecanismo explicativo propio de la actitud intencional<sup>105</sup>. Según esto, la atribución de contenidos intencionales a un organismo depende de una asunción optimizadora respecto de sus

---

<sup>104</sup> Dennett (1987a, p. 52).

<sup>105</sup> Como vimos en el apartado 2.2 (p. 42), Fodor mismo proporciona un argumento anti-evolucionista en contra de esta tesis.

estados mentales, asunción que Dennett formula en términos de la apelación a principios de racionalidad. La consecuencia inmediata de esto es que el vocabulario intencional cobra un carácter idealizador, aproximado, impreciso<sup>106</sup>. Un punto ulterior que destacábamos es que la misma eficacia de la actitud intencional es explicable por este carácter, en la medida en que permite al intérprete obviar una cantidad de información “visible” desde las actitudes física y de diseño.

Ahora puede verse cómo el enfoque de la psicología popular desde la distinción creencia / opinión puede complementar la visión general del discurso intencionalista en tanto que resultado de la adopción de la actitud intencional. El hecho de que dicho discurso sea entendido como inherentemente impreciso (dada la naturaleza del mecanismo explicativo que lo sostiene) y, por otra parte, el hecho de que sea entendido fundamentalmente en términos de su contribución pragmática en la regulación del comportamiento vienen a coincidir en una explicación de por qué la psicología popular funciona. Más exactamente, la tesis dennettiana de que el discurso psicológico popular funciona *porque* es impreciso sólo cobra sentido en el contexto de una interpretación pragmatista de su rol e importancia. Y es en este punto donde, a mi parecer, puede dirimirse la disputa entre Dennett y Fodor a este respecto.

De hecho, al margen de cualquier otra consideración que pueda hacerse respecto del discurso psicológico popular en cuanto tal, hay una diferencia básica en el enfoque de ambos autores, que es la que hay entre interpretar un discurso mayormente por su aspecto descriptivo (esto es, esencialmente, como un discurso cuyos términos refieren a aspectos del mundo) e interpretarlo por su aspecto pragmático (esto es, esencialmente, como un discurso que responde a determinadas motivaciones y genera determinadas consecuencias en un contexto social). Esta distinción es sumamente importante para abordar la disputa entre las diferentes concepciones de la psicología de sentido común aceptadas por los autores.

La crítica estándar a la concepción pragmatista, no sólo por parte de Fodor sino de quienquiera que acepte una interpretación realista del discurso psicológico popular, puede expresarse así: “¿si la psicología popular no es verdadera, entonces por qué funciona tan bien?”. Nótese lo que la crítica da por supuesto: esto es, que hay una conexión directa entre la verdad de una teoría y su utilidad. Hay que subrayar que

---

<sup>106</sup> La consecuencia mediata es que el vocabulario de la psicología de sentido común no puede funcionar, sin más, como *modelo* para una psicología científica de nivel sub-personal. Profundizaré sobre esto en el apartado 4.4.

Dennett en ningún modo negaría esto. Más aún, hay un sentido en que aceptaría plenamente que la psicología popular es verdadera. Sin embargo, dado su enfoque, éste no es el mismo sentido en que lo aceptaría Fodor.

El enfoque descriptivista de Fodor se sustenta en una noción general de la verdad de un discurso mucho más restringida de la que Dennett podría aceptar. Para Fodor, la evaluación de la verdad de un discurso depende de modo directo de la medida en que sus términos refieran a entidades, propiedades y procesos en el mundo. Si el discurso produce de modo fiable predicciones y explicaciones a través de la postulación de estas entidades, propiedades y procesos, entonces no sólo el discurso es verdadero, sino que de su verdad puede derivarse un compromiso más o menos fuerte con lo que sea que postule.

Es claro que este enfoque para dar cuenta de por qué la psicología popular funciona encuentra asiento en la concepción intelectualista de Fodor. Más aún, mi idea es que la justificación fodoriana de muchos de los rasgos que esta concepción atribuye al discurso psicológico popular está fuertemente influenciada por un interés programático respecto de la elaboración de una psicología científica (verdadera). Quiero intentar mostrar por qué este interés resulta ser altamente restrictivo respecto del fenómeno complejo de la psicología de sentido común.

Recordemos brevemente los aspectos centrales de la concepción intelectualista. La idea básica, *contra* Dennett, es entender el discurso mentalista como una suerte de teoría. Esto tiene dos aspectos: en primer lugar, el discurso mentalista está estructurado de modo similar a las construcciones teóricas propias de las ciencias especiales (aunque con diferentes grados de precisión y rigidez en las inferencias), lo que incluye especialmente el recurso a entidades no observables; en segundo lugar, las generalizaciones que implícitamente realizamos en el discurso cotidiano son del mismo tipo que las generalizaciones explícitas de las teorías científicas.

Creo que hay una crítica general que puede hacerse a este modo de presentar la psicología popular. En primer lugar, hay que señalar que las características que se le atribuyen responden, como mencioné más arriba, a un enfoque determinado para el análisis de un discurso (esto es, el que se concentra en su aspecto meramente descriptivo); este enfoque es claramente favorecido por el interés fodoriano (que, como tal, es plenamente justificable) en determinar el papel que el discurso psicológico popular puede jugar en la conformación de una psicología científica. Ahora bien, el punto del argumento es que el enfoque descriptivo, por más justificado que esté de

manera independiente, es restrictivo, esto es, no proporciona una perspectiva adecuada para dar plenamente cuenta de su objeto: mientras dar primacía al enfoque pragmático deja lugar para explicar las características de la psicología popular en tanto que discurso que refiere al mundo, dar primacía al enfoque descriptivo *restringe* el foco de análisis a sólo esas características.

La concepción pragmatista, armada, según he sostenido, de la distinción entre creencia y opinión y, con ella, de la distinción entre habilidad e ideología, puede circunscribir los aspectos descriptivos de la psicología popular a uno de los términos distinguidos: esto es, el conjunto, implícito o explícito, de las creencias (en sentido técnico, de las opiniones) que *refieren* a la psicología popular en su sentido más básico. Estas creencias podrán variar de sujeto en sujeto, de cultura en cultura, y aún así no afectarán mayormente el uso y la efectividad del vocabulario mentalista. De este modo resulta claro que la relación implícita en la pregunta del realista, “si es falso, ¿por qué funciona?”, entre la verdad y la utilidad de un discurso responde a un enfoque (el enfoque descriptivo) que no hace justicia a la realidad concreta de la psicología de sentido común. Más aún, enfocar la psicología popular primeramente desde su aspecto pragmático todavía permite formular la pregunta (en la que Fodor se centra) por el posible aporte del discurso intencionalista a una psicología científica y a otras disciplinas relacionadas<sup>107</sup>.

Otro aspecto negativo de la concepción intelectualista, también derivado de no incorporar las distinciones dennettianas, es que asume (aunque no defiende explícitamente) un carácter unitario u homogéneo en el discurso psicológico popular, y por analogía con la teorización científica, que resulta difícil de defender. La concepción pragmatista es más compatible con el hecho de que, en última instancia, dicho discurso es un producto popular, influenciado como tal por factores diversos, absolutamente cambiantes, arbitrarios, no controlables.

En primer lugar, puede darse una mejor explicación de este hecho si se acepta que el núcleo de la psicología popular es una habilidad para hacer predicciones y tener expectativas sobre el comportamiento ajeno. En segundo lugar, respecto de su aspecto más descriptivo, quiero recordar el carácter que llamé “liberal” (p. 62) del intercambio de opiniones, que puede dar cuenta del uso descuidado que cotidianamente hacemos de las entidades y procesos postulados por el discurso mentalista. La imagen que nos queda

---

<sup>107</sup> En 4.4, de hecho, me concentraré en la postura defendida por Dennett a este respecto.

de la psicología popular es la de un fenómeno heterogéneo, diverso y cambiante, pero, como también destacaba en p. 66, esta diversidad es compatible con la idea de que ambos aspectos de la psicología (así como los mismos conceptos de creencia y opinión) están fuertemente integrados, por más que puedan aislarse teóricamente.

Por otra parte, la concepción fodoriana, con su énfasis puesto fuertemente en la contribución del discurso popular al desarrollo teórico científico, plantea la cuestión intuitivamente poco feliz de la *necesidad* de legitimar los supuestos de la psicología de sentido común, legitimación que para Fodor consistiría en la especificación de condiciones naturalistas (no intencionales o semánticas) para que un estado físico tenga contenido intencional<sup>108</sup>. Esta manera de enfocar la contribución de la psicología popular no la vería sin más como teoría, sino como una teoría verdadera en espera de *confirmación y reducción*. El fracaso en esta tarea implicaría el fracaso de la teoría.

Es a primera vista dudoso, sin embargo, que de este eventual fracaso se seguirían consecuencias terribles en el modo en que comúnmente nos entendemos y actuamos en sociedad<sup>109</sup>. Con palabras de Fodor:

[S]i la psicología intencional del sentido común fuera realmente a colapsar, ésta sería, más allá de toda comparación, la catástrofe intelectual más grande en la historia de nuestra especie. Si estamos equivocados con relación a la mente, entonces éste es el peor error que hemos cometido acerca de nada<sup>110</sup>.

El “colapso” de la psicología intencional equivaldría a la demostración de que, interpretada como teoría, ésta no fuera *literalmente* verdadera. Aquí resurge la noción restringida de verdad, mencionada anteriormente, que el enfoque descriptivo asumido por Fodor favorece, y que sustenta la “teoría de la catástrofe”. Es exactamente en este mismo sentido restringido, por otro lado, que para Dennett, y a la luz de su enfoque pragmatista para la explicación del éxito del discurso psicológico popular y de las características que le atribuye, no podría hablarse estrictamente de la verdad de dicho discurso (y, *a fortiori*, es por esto mismo que no se seguirían consecuencias catastróficas en ausencia de una teoría naturalizada del contenido<sup>111</sup>).

---

<sup>108</sup> Este constituye, básicamente, el problema de la naturalización del contenido, presentado en el apartado 2.1 (pp. 36-37).

<sup>109</sup> Stephen Stich (1996, pp. 168-171) denomina algo irónicamente “Teoría de la Catástrofe” la idea fodoriana de que, en ausencia de una teoría que naturalice lo mental, se seguiría el irrealismo intencional y, por lo tanto, las categorías intencionales tendrían que ser abandonadas.

<sup>110</sup> Fodor (1987, pp. 14-15).

<sup>111</sup> No sólo esto, sino que, dado el enfoque dennettiano sobre el contenido mental, también habría razones en contra de la misma coherencia y viabilidad de una tal teoría. Avanzaré sobre esto en el apartado 4.3.

Además de lo ya explicitado en relación con su adopción de un enfoque descriptivo, hay otros dos sentidos por los que creo pueda acusarse a la concepción intelectualista de ser restrictiva: el primero respecto del innatismo de la psicología popular y el segundo respecto de su posible extensión a formas no humanas de atribución intencional. En cuanto a este segundo punto, aunque Fodor nunca critica explícitamente la adecuación de hablar de “psicología intencional” para el caso de organismos no humanos, tampoco la defiende o siquiera toma en consideración el asunto<sup>112</sup>. El punto está en que su concepción de la psicología intencional sencillamente no deja lugar para el planteo del problema en este contexto extendido. De un modo muy claro y directo, modelar la psicología intencional principalmente por su afinidad general con nuestros constructos teóricos explícitos y sofisticados deja poca esperanza para investigar su aplicabilidad en seres evidentemente incapaces de manipular y mucho menos elaborar algo que siquiera se aproxime a tales constructos.

Es claro cómo la diferencia entre creencias y opiniones entra plenamente en juego aquí. La analogía con el formato de una teoría científica está subliminalmente incorporando una relación necesaria entre la idea de una psicología intencional y la competencia lingüística, entendida exclusivamente como la competencia característicamente humana para manipular un lenguaje articulado. La concepción intelectualista, entonces, es restrictiva en este punto, en un sentido directo y general: asume un compromiso por el cual se entiende que, como hecho general, la psicología popular implica competencia lingüística. Esto tiene obvias consecuencias respecto de la aplicabilidad de una noción así construida de psicología popular.

Ahora bien, más allá de este carácter restrictivo general, la concepción intelectualista se encontraría en dificultades concretas si se le pidiera de dar cuenta de los resultados de la creciente investigación en etología cognitiva, inaugurada por Premack y Woodruff (1978), que trabaja explícitamente el problema de la teoría de la mente en primates no humanos. No poder dar cuenta de los resultados de esta línea de investigación es, por lo menos, un problema para Fodor, si no un aspecto negativo importante de su concepción.

Por otra parte, justamente en este punto es donde la concepción pragmatista de Dennett, con las distinciones que incorpora, se fortalece. Sin ir más lejos, el núcleo de dicha concepción, esto es, el mecanismo de la actitud intencional, ha sido en general

---

<sup>112</sup> Este es un caso puntual de una actitud metodológica general de Fodor, a la que me referiré en la discusión de su propuesta de una “psicología de las actitudes proposicionales” (pp. 94-95).

acogido en el campo de la etología cognitiva como una herramienta teórica extremadamente útil y ha de hecho ampliado notablemente el foco de las discusiones<sup>113</sup>. El uso que en este campo se ha hecho de la noción de “actitud intencional” trasparenta un hecho a tener en cuenta aquí: la concepción pragmatista incorpora, y ayuda a configurar, una importante distinción entre, por un lado, ser susceptible de descripción, explicación o predicción desde la actitud intencional y, por otro lado, ser susceptible de adoptarla.

El problema de la *capacidad* para la adopción de la estrategia intencional es, en cierta medida, una reformulación de la pregunta por la aplicabilidad de la noción de “psicología intencional”, pregunta que se vuelve de interés una vez que se consideren casos de comportamiento social sofisticado en animales no humanos. Si para dar cuenta de tales casos entre dos animales, *a* y *b*, fuera necesario recurrir no sólo al vocabulario mentalista de primer orden (“*a* cree que *P*”) sino al de segundo orden, esto es, y valga la redundancia, atribuir atribuciones intencionales (“*a* cree que *b* cree que *P*”), entonces, de acuerdo al enfoque dennettiano, estaríamos justificados en hablar de estos animales como “psicólogos populares”. De hecho, Dennett (1991d) utiliza dos expresiones que reflejan esta postura a la luz de la distinción entre creencias y opiniones: habla, en particular, de los chimpancés como “psicólogos populares individuales” (*individual folk psychologists*), contrastándolos con nosotros, seres humanos, que seríamos “psicólogos populares comunales” (*communal folk psychologists*), dotados de la posibilidad que nos brinda el lenguaje de explicar y pedir explicaciones de las razones de nuestros comportamientos.

Además, nuevamente desde la perspectiva de la distinción creencia / opinión, y en contraste con el carácter restrictivo de la concepción intelectualista, queda claro que el espectro de los tipos de atribuciones intencionales que la concepción pragmatista puede albergar es más abarcativo, y la misma noción de atribución intencional es consecuentemente más flexible. El enfoque amplio de la concepción dennettiana encuentra expresión en las atribuciones inespecíficas de la etología cognitiva: esto es, atribuciones no mediadas por un lenguaje natural articulado (en el sentido de que no son realizadas a través de los medios lingüísticos por los que nosotros comúnmente atribuimos contenidos mentales) ni respaldadas<sup>114</sup> por un lenguaje natural articulado (en

---

<sup>113</sup> Ver el artículo de Robert Seyfarth y Dorothy Cheney, “Dennett’s Contribution to Research on the Animal Mind”, en Brook y Ross (eds.), (2002).

<sup>114</sup> En el mismo sentido en que uso el término en p. 72.

el sentido de que las atribuciones no son implícita ni explícitamente acompañadas por una pre-concepción aproximada sobre lo que se atribuye o, lo que es lo mismo, por un conjunto difuso de opiniones sobre lo que se atribuye).

A la vez, hay que mencionar la importancia que este enfoque amplio puede tener para una perspectiva específicamente evolucionista respecto de la psicología popular, del mismo modo como, por otro lado, el enfoque restrictivo fodoriano cierra el camino para dicha perspectiva<sup>115</sup>: este último, como vimos, de hecho no proporciona herramientas para poder extender el análisis hasta lo que, desde una perspectiva evolutiva, podrían considerarse formas más rudimentarias de atribución intencional, reconocidas y estudiadas en los organismos evolutivamente más cercanos a nosotros. Algunos de los rasgos comunes entre estas formas más rudimentarias y nuestras atribuciones intencionales sofisticadas podrían aislarse y considerarse la base para una investigación evolucionista de los orígenes de la psicología popular (humana). Indudablemente, los rasgos relacionados con la capacidad general para la interacción social serán los más relevantes y, en este sentido, creo que los actuales estudios sobre empatía y cognición social en primates son exponentes ilustres del tipo de investigación a la que hago referencia<sup>116</sup>.

Pasando ahora al problema del innatismo, es interesante notar que, al igual que para la cuestión de la verdad del discurso psicológico de sentido común (p. 74), las concepciones pragmatista e intelectualista acordarían en aceptar algún grado de innatismo en la psicología popular; pero, así como las concepciones disienten respecto del sentido en que la psicología sería verdadera, disienten claramente también respecto del sentido en que aceptarían que es innata. En lo que refiere a este punto, la concepción intelectualista es restrictiva en tanto que, dado su enfoque restrictivo general señalado antes, adopta una versión rígida de innatismo que implica entre otras cosas un compromiso fuerte con la idea de que hay conceptos innatos.

Aunque ésta pudiera ser una posición defendible, y aunque de hecho el atomismo conceptual<sup>117</sup> de Fodor incorpora una defensa de un innatismo de este tipo, muchos autores acuerdan en que esta defensa comporta una carga explicativa muy

---

<sup>115</sup> El problema, como antes (pp. 74-75), no es que la concepción intelectualista no tome en consideración o simplemente no considere relevante cierta perspectiva de estudio sino que, tal como la hemos analizado, no deje lugar para ella.

<sup>116</sup> Entre otros, en Frans de Waal (1996), se resumen muchos de estos estudios.

<sup>117</sup> La exposición más completa de esta postura, junto con un correspondiente y vigoroso ataque a la llamada “semántica del rol inferencial” (y también a su contraparte en Inteligencia Artificial, la “semántica procedural”) se encuentra en *Concepts* de 1998. Los detalles de la teoría, así como sus disputas con los opositores, no son relevantes para la discusión.

pesada. Mi intención aquí no es la de evaluar el atomismo conceptual, sino que para mis fines es suficiente indicar que, si a una concepción del discurso intencionalista que lo interprete esencialmente como una teoría se le agrega la tesis de que, con palabras de Fodor, “el *conocimiento* de la psicología de sentido común del *Homo sapiens*”<sup>118</sup> es innato, entonces la idea resultante de “innatismo” y de “conocimiento innato” arrastrará necesariamente connotaciones y limitaciones propias de dicha particular interpretación.

Un corolario de lo anterior es que los argumentos propuestos por Fodor para respaldar su tesis de que la psicología popular es innata no apuntan a defender específicamente la versión fuerte de innatismo que la concepción intelectualista requiere. Los argumentos en cuestión son los siguientes: (1) la explicación en términos intencionales parece ser universalmente aceptada por las diferentes culturas; (2) buena parte del aparato explicativo mentalista ya es operativo en el niño a una edad muy temprana; (3) no hay hipótesis rivales a la opción innatista, lo cual reflejaría la dificultad de dar cuenta de la adquisición del aparato mentalista sin apelar a ella. Sobre todo en este último caso, queda claro que no se contempla la posibilidad de dar versiones alternativas o simplemente diferentes de la tesis innatista. El recorte a la idea de un conocimiento innato de la psicología popular, que como veíamos es exigido por las restricciones que implican entenderla como una especie de teoría, es asumido por los argumentos, y es por lo tanto necesario apelar a una teoría independiente, con un pesado bagaje explicativo independiente (esto es, el atomismo conceptual fodoriano), para definir y volver científicamente admisible un innatismo conceptual fuerte como el requerido.

Por otra parte, la concepción pragmatista de la psicología de sentido común puede dar lugar a la tesis de que ésta, o más bien algunos de sus aspectos, puedan ser innatos. Esta versión de innatismo sin embargo en absoluto incorpora la necesidad de aceptar la existencia de conceptos innatos. La versión pragmatista de la idea de que la psicología popular es innata es justamente aquella en virtud de la cual estaría, como afirmábamos más arriba, justificado extender la concepción de la psicología intencional a algunos animales no humanos.

Haciéndonos nuevamente eco del carácter disyuntivo de la concepción pragmatista, sus aspectos más claramente disposicionales, concebidos en términos de habilidades, son los únicos a los que debe apelarse para configurar la tesis del innatismo

---

<sup>118</sup> Fodor (1987, p. 189). El subrayado es mío.

de la psicología popular. Dada la distinción entre creencias y opiniones, y con ella la distinción entre habilidad e ideología, no es en ningún caso necesario (ni en el caso de los animales no lingüísticos ni en los humanos) apelar a la noción intuitivamente poco viable y explicativamente muy costosa de “concepto innato”.

Ya concluyendo mi análisis comparativo de las concepciones pragmatista e intelectualista de la psicología popular, quiero destacar un punto de suma importancia para el resto de las discusiones a tratar y que es, a mi parecer, una consecuencia directa de la respectiva aceptación, por parte de Dennett y Fodor, de dichas concepciones. A la vez, es la principal fuente de crítica que la distinción entre creencias y opiniones aporta, crítica que se ramifica en el fondo hacia todas las disputas entre nuestros autores.

Como vimos, la concepción pragmatista permite distinguir, apoyándose en los conceptos de creencia y opinión, dos variantes o aspectos en la psicología popular con la que nosotros, seres humanos, nos enfrentamos al mundo. Ahora, el vocabulario mentalista que usamos día tras día incorpora inevitablemente ambos aspectos, y este hecho tiene una relevancia metodológica que, como espero poder mostrar en adelante, se extiende a cualquier punto de una teoría de la mente. El aspecto lingüístico, culturalmente afectado, de la psicología popular asume las características propias de las opiniones dennettianas y, entre ellas, destacan algunas por sus implicaciones metodológicas: la “liberalidad” en el empleo de la psicología, lo que incluye el hecho de que no requiere un manejo preciso y riguroso de sus términos; su carácter derivado, lo que incluye su relativa distancia respecto de la conducta de la que pretende dar cuenta; y el hecho relacionado de que es flexible y sujeta a variaciones culturales.

Si a estas características se las interpreta en conjunto en el contexto de la tesis central dennettiana, según la cual la atribución intencional se proyecta a través de una ineliminable asunción de racionalidad, el carácter inherentemente impreciso, aproximado, abstracto y optimizador que de este modo el vocabulario mentalista asume se redimensiona y profundiza. El lenguaje mentalista, inevitablemente impregnado de ideología, llega a ser concebido como una construcción, un artefacto útil, regido fundamentalmente por presiones sociales contextuales, y cuya adecuación y coherencia interna, la validez de sus generalizaciones, así como el valor referencial de sus términos son todos aspectos subsidiarios respecto del rol pragmático que cumple. Dadas las mencionadas características, hay un sentido importante en que, aún admitiendo plenamente su crucial rol pragmático, no estaríamos por esto habilitados para aceptar confiadamente aquellos aspectos. Por lo tanto, en el caso de que nuestro interés sea de

índole científico, esto es, en el caso de que sí nos interese la confirmación y el examen de dichos aspectos, el lenguaje mentalista deberá ser tomado bajo importantes restricciones y con un cuidado especial: ésta es, en términos generales, la enseñanza metodológica que deja la concepción pragmatista<sup>119</sup>.

La concepción intelectualista fodoriana, y con ella la teoría representacional de la mente, claramente no asume este punto de vista, este llamado de atención metodológico. De este modo, la teoría queda expuesta, desde la perspectiva dennettiana, al aducido carácter engañoso del discurso intencionalista. Puede verse aún más claramente que esto es así si tenemos en mente el carácter fuertemente científicista de la concepción fodoriana del discurso psicológico popular, enfocado en su totalidad, tal como sostuve, desde una perspectiva descriptiva. Un enfoque de este tipo, junto con el interés fodoriano en la contribución directa del discurso mentalista a una psicología científica, tiende a favorecer implicaciones metodológicas contrarias a las recién explicitadas.

La expresión concreta de estas implicaciones es un marco metodológico para el estudio de los fenómenos intencionales que, hace más de sesenta años, es dado por supuesto en la gran parte de las discusiones en la filosofía contemporánea de la mente y, de hecho, es mayoritariamente considerado el modelo estándar de la investigación filosófica sobre los estados intencionales. (No es, en cambio, un modelo al que forzosamente se ciñan las investigaciones en psicología cognitiva). Este modelo es el del análisis en términos de *actitudes proposicionales*.

## **2. Las Actitudes Proposicionales como Modelo de Análisis Filosófico**

La noción de “actitud proposicional” fue propuesta primeramente por Bertrand Russell (1940) y refiere a los estados intencionales en general, proporcionando un formato determinado para ellos. En lo que sigue voy a exponer este formato, especialmente del modo en que es asumido por Fodor, y voy a presentar y discutir las críticas que Dennett le hace (o podría hacerle)<sup>120</sup>, siempre desde la perspectiva del par creencia / opinión. Mi

---

<sup>119</sup> Esta es, reinterpretada, la misma “enseñanza” que destacaba en p. 67, para la distinción creencia / opinión.

<sup>120</sup> Como para el caso de las concepciones de la psicología de sentido común, la serie de críticas que voy a presentar no son explícitamente expuestas por Dennett, sino que pueden configurarse desde la particular interpretación que vengo defendiendo de la teoría de los sistemas intencionales y de su enfoque

sospecha, en este sentido, es que hay un vínculo importante entre la adopción del enfoque pragmatista hacia la psicología popular y el rechazo del análisis en términos de actitudes proposicionales, así como lo hay entre la adopción del enfoque descriptivo y la aceptación de dicho análisis: más expresamente, intentaré mostrar que, considerando la advertencia metodológica dennettiana antes expuesta, el modelo fodoriano de las actitudes proposicionales puede considerarse en algún grado como una consecuencia directa de, y cuya justificación se basa en, su concepción intelectualista, concepción que se apoya, según espero haber mostrado, en una visión inadecuada del papel del lenguaje en la psicología popular. Del mismo modo, la distinción entre creencias y opiniones, en tanto que herramienta argumentativa, podrá interpretarse como yendo directamente en contra del análisis en términos de actitudes proposicionales. Hablaré, en este sentido, de la “tesis de las actitudes proposicionales” (TAP), como aquella que sostiene que la noción de actitud proposicional aporta el modelo apropiado para el abordaje filosófico de los problemas concernientes al contenido mental y los estados intencionales.

El formato de las actitudes proposicionales es un supuesto profundamente arraigado en la teoría representacional de la mente, tanto que en ocasiones Fodor habla de psicología de las actitudes proposicionales y de psicología de sentido común indiferentemente. Puede decirse que el análisis en términos de actitudes proposicionales es el núcleo metodológico de la teoría<sup>121</sup>: su importancia radica en que aporta un modelo estándar para la atribución intencional. El modelo es el siguiente: [sujeto de atribución – tipo de actitud – objeto proposicional (o contenido proposicional)], y se ejemplifica en oraciones como “*a* cree que *P*”, “*b* desea que *Q*”, “*c* teme que *R*”. Nótese que de hecho no hablamos de este modo comúnmente o, si lo hacemos, sólo raramente: el modelo no se propone reflejar el modo en que hablamos, sino que pretende aportar una estructura básica idealizada para toda atribución intencional.

Las ventajas claras de adoptar este enfoque son: (1) que obtenemos un modelo fijo para la individuación de los estados intencionales y (2) que estandarizamos una manera clara de hablar sobre los estados intencionales<sup>122</sup>. El modelo admite tres elementos a combinar: uno para el sujeto de la atribución, otro para el tipo de actitud en

---

argumentativo general. Dennett expone, sin embargo, una serie de críticas al modelo de las actitudes proposicionales en el conocido “Beyond Belief” (1982b).

<sup>121</sup> En “Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology”, Fodor (1981, p. 226) llega a decir que cualquier teoría que tome la línea explicativa de las actitudes proposicionales (esto es, cualquier teoría que acepte TAP) es, sin más, una versión de la teoría representacional de la mente.

<sup>122</sup> Las variantes del modelo de hecho tienen sus respectivos análogos gramaticales: sujeto para el objeto de atribución, verbo para el tipo de actitud y complemento para el objeto proposicional.

la que se encuentra y otro para el contenido proposicional que ésta especifique (esto es, diferentes sujetos pueden relacionarse intencionalmente de diferentes modos con diferentes contenidos proposicionales); y una atribución intencional será tal (y, de este modo, un estado mental será individuado) si las tres variantes se fijan.

Fodor proporciona la especificación y defensa más completa de TAP en el artículo temprano acordemente llamado “Propositional Attitudes” (“Las Actitudes Proposicionales”) (1978). En ese artículo, Fodor acepta y defiende de modo explícito la interpretación de las creencias, deseos y demás estados intencionales como actitudes proposicionales: su estrategia para esto es presentar cinco condiciones *a priori* a ser satisfechas para obtener lo que llama una “teoría de las actitudes proposicionales” y argumentar luego que la psicología cognitiva empírica de hecho trabaja sobre una noción de actitud proposicional como la que dichas condiciones especificarían. Quiero mostrar que el modo en que por esta vía se justifica la pertinencia de la interpretación de los estados intencionales como actitudes proposicionales refleja claramente, y en diversos puntos, un paso argumentativo implícito que sólo se sustenta en el marco de la concepción intelectualista del lenguaje psicológico.

En “Beyond Belief” (1982b)<sup>123</sup>, Dennett distingue tres características generales y poco compatibles de las proposiciones en la literatura filosófica: las proposiciones entendidas o bien como entidades parecidas a oraciones (es decir, estructuradas de acuerdo a una sintaxis), o bien como grupos de mundos posibles (esto es: dos proposiciones son equivalentes si son verdaderas para el mismo grupo de mundos posibles), o bien como disposiciones de objetos y propiedades en el mundo (esto es: dos proposiciones son equivalentes si atribuyen las mismas propiedades a los mismos objetos). Fodor adheriría enteramente a la primera de estas caracterizaciones: lo que sigue es un intento de mostrar por qué esta toma de posición no es filosóficamente ingenua.

La primera condición propuesta por Fodor para una teoría de las actitudes proposicionales es que éstas deben ser analizadas como *relaciones*. Ahora, Fodor entiende estas relaciones en un sentido en ningún modo laxo, esto es, como relaciones cuyos términos son objetos o conjuntos de objetos determinados y aislables:

En particular, el verbo en una oración como “Juan cree que está lloviendo” expresa una relación entre Juan y *algo*, y un ejemplar de

---

<sup>123</sup> En Dennett (1987a, pp. 114-116 aprox.).

esa oración es verdadero si y sólo si Juan está en la relación-de-creencia con *ese algo*<sup>124</sup>.

Los únicos dos argumentos que da Fodor para defender esta primera condición son los siguientes<sup>125</sup>: (1) un argumento según el cual “‘cree’ parece ser una relación binaria, y sería bueno que nuestra teoría de la creencia nos permitiera salvar las apariencias”<sup>126</sup> y (2) un argumento que pretende derivar de la posibilidad de hacer una generalización existencial (de “Juan cree que está lloviendo” a “Hay algo que Juan cree”, por ejemplo) un “indicio” (!) de compromiso ontológico. En ambos casos, parece evidente que lo que Fodor está haciendo es intentar derivar algún tipo de “relacionalidad” fuerte (es decir, en algún grado, empírica) a partir de consideraciones sobre las expresiones de actitudes proposicionales. Pero continuemos.

Las siguientes dos condiciones puestas por Fodor son las que llama “condición de Vendler” y “condición de Frege”. La primera exige que una teoría de las actitudes proposicionales explique el paralelismo que hay entre los verbos de actitudes proposicionales y los verbos que corresponden a decir. El hecho del que se pretende dar una explicación aquí es que “[e]n general, las cosas que podemos decir que *creemos* (deseamos, confiamos, lamentamos, etcétera) son las mismas cosas de las que podemos decir que *decimos* (aseveramos, enunciamos, etcétera)”<sup>127</sup>. Por otra parte, la condición de Frege exige que se dé cuenta de la opacidad referencial de las actitudes intencionales, el fenómeno lógico por el cual principios de inferencia como la mencionada generalización existencial y la substitutividad de idénticos no se aplican en casos de fórmulas lingüísticas anteceditas por un verbo de actitud proposicional (mientras sí se aplican para las mismas fórmulas no anteceditas por el verbo). Para el caso de ambas condiciones, se considera necesaria la explicación de fenómenos que caen en un contexto claramente lingüístico.

Una cuarta condición para una teoría de las actitudes proposicionales, posiblemente la más importante, es la llamada “condición de Aristóteles”: ésta es la idea de que los objetos de las actitudes proposicionales tienen *forma lógica*. El punto central es que esta forma lógica haría no arbitrarias las interacciones causales entre los estados mentales, interacciones que a la vez constituirían los procesos mentales que resultan

---

<sup>124</sup> Fodor (1978, p. 174). El subrayado es mío.

<sup>125</sup> Habría un tercer argumento, pero no es nada más que el reclamo de que hasta el momento no se han dado explicaciones más convincentes.

<sup>126</sup> Fodor (1978, p. 175).

<sup>127</sup> Fodor (1978, p. 178).

finalmente en las conductas inteligentes de los organismos. Por último, la quinta condición que Fodor propone refiere a la viabilidad empírica de la teoría: ésta debería poder asociarse a explicaciones empíricas de los procesos mentales, y ser de este modo compatible con los datos experimentales sobre las que éstas se apoyen. Esta última condición es claramente distinta de todas las demás y no será tratada aquí.

Lo que las cuatro primeras condiciones propuestas por Fodor tienen en común es que todas de alguna forma están enmarcadas dentro de un contexto lingüístico o bien hacen referencia a algún hecho lingüístico particular. El defecto mayor de la interpretación de los estados intencionales como actitudes proposicionales es que puede llevar a quienes la adopten a utilizar criterios meramente lingüísticos respecto de cómo usamos el discurso mentalista cotidiano en el intento de determinar restricciones teóricas para una psicología científica, o proporcionar un modelo para ella, o incluso directamente para probar hechos psicológicos sustantivos respecto de la implementación de los estados mentales a nivel sub-personal.

En ausencia de una distinción entre creencias y opiniones, los elementos teóricos propiamente lingüísticos que Fodor introduce para configurar la idea de una teoría de las actitudes proposicionales, entendida como una teoría de la intencionalidad para los estados psicológicos, pasan desapercibidos y de este modo no juegan un papel en tanto que elementos lingüísticos. Es importante recordar en este sentido que las condiciones propuestas son consideradas *a priori*, lo cual evidencia aún más el hecho de que las características introducidas son dadas por supuestas o aceptadas, incluso, como conceptualmente evidentes; especificando el carácter *a priori* de sus condiciones, Fodor dice: [a]lgunos de los puntos que defenderé son, supongo, estrictamente conceptuales, pero otros son meramente *autoevidentes*”<sup>128</sup>. Concentrémonos ahora en la primera condición, por la cual las actitudes proposicionales han de entenderse como relaciones.

¿Qué significa esto? La condición no está especificando un formato que autoriza a título heurístico una cierta forma de hablar sobre los estados mentales. Una cierta forma de hablar sobre los estados es el *punto de partida* a partir del cual la condición se especifica, esto es, el hecho de que formulamos atribuciones intencionales de un cierto modo. Y lo que, apelando únicamente a este hecho, la condición requiere es que se entiendan los estados mentales (y no nuestras atribuciones lingüísticas de ellos) efectivamente como algún tipo de relación instanciada de alguna manera determinada.

---

<sup>128</sup> Fodor (1978, pp. 173-4). El subrayado es mío.

Ahora bien, “cree”, “desea” y demás actitudes denotan relaciones binarias exactamente en el mismo sentido en que lo hacen muchas otras construcciones en castellano o en cualquier idioma (“Dr. Jeckyll es Mr. Hyde”, “el chico come carne”, etc.), y únicamente del hecho de que puedan verse las expresiones de actitudes proposicionales como la unión de dos elementos lingüísticos (el sujeto de la atribución, por un lado, y el contenido proposicional, por otro) no se sigue que estemos justificados para aislar uno de dichos elementos (el contenido proposicional) y considerarlo como tal un elemento físicamente instanciado, integrante de algún fenómeno mental particular. Como dije, considero que “creer”, en un sentido totalmente laxo, expresa una relación. Ahora, el problema es determinar el papel teórico que hacemos jugar a las actitudes proposicionales si aceptamos su estatus adquirido de relaciones y, consecuentemente, determinar cuál es la definición precisa de dicho estatus y los compromisos que ésta comportaría.

El mismo Fodor, al especificar la primera condición, trae a colación la cuestión de la carga (en este caso, la carga ontológica) que el lenguaje nos lleva a aceptar. Aborda el problema apoyándose en un ejemplo del mismo Dennett (1969), el caso de “la voz de María”, expresión que parecería llevarnos a aceptar la existencia de entidades-como-voces que en este caso María poseería. La razón que da Fodor para justificar, para este ejemplo en particular, el llamado de atención dennettiano por el que consideraríamos la gramática del lenguaje como engañosa es que podemos transformar sin pérdida de significado la expresión en cuestión en otra expresión ontológicamente menos comprometida; Fodor, por ejemplo, sugiere transformar en este sentido la expresión “María tiene una buena voz” en la expresión “María canta bien”.

Ahora bien, en primer lugar, es dudoso que esta propuesta transformación sea un método confiable y aplicable ubicuamente para cualquier expresión de cualquier lenguaje natural: el mismo caso de “María canta bien” parece no ser semánticamente homologable a su análogo ontológicamente cargado (podría objetarse que “María canta bien” es más restringido que “María tiene una buena voz” ya que hace referencia al acto específico de cantar: una cosa es calificar de cierta forma la voz de María, otra diferente es referirse al hecho de que ella cante bien); en segundo lugar, hay que tener en cuenta el modo desigual en que Fodor y Dennett ponen en juego su parsimonia metodológica: Fodor lo hace a modo de recurso circunstancial para casos en que el lenguaje nos conduciría a mantener ideas claramente contraintuitivas, mientras que Dennett adopta el

criterio como un principio general para abordar todo tipo de casos, por más familiares que nos parezcan (o, aún, especialmente para éstos).

De hecho, como señalé, la idea dennettiana de una “ciencia popular” o de una ideología impregnada en nuestro vocabulario mentalista y en el modo en que lo empleamos viene justamente a respaldar esta segunda manera de adoptar el cuidado metodológico respecto de las implicaciones lingüísticas, para el caso del discurso psicológico de sentido común. En este sentido, no importa en absoluto el hecho de que una implicación lingüística nos parezca intuitivamente aceptable o no, por más fuerte que esta convicción sea.

Por otro lado, el modo en que Fodor adopta este cuidado depende enteramente de las ideas intuitivas generadas por nuestro constante uso del lenguaje, y del grado en que estas ideas nos sean más o menos familiares. De esta manera, dejados de lado los casos que a Fodor no molestan, los dos argumentos que presenta para la primera condición pueden asentarse, como lo hacen, en la analogía con algunas expresiones lingüísticas corrientes. Lo mismo vale para las condiciones de Vendler y de Frege. Esta última demanda la explicación del fenómeno lógico-semántico de la opacidad referencial y la condición de Vendler es, de hecho, un requisito explícito de que la mencionada analogía sea explicada: se pide que una teoría psicológica dé cuenta del hecho absolutamente lingüístico de que las mismas cosas de las que *decimos* que creemos son las mismas de las que *decimos* que decimos.

El cuidado metodológico respecto de nuestras expresiones de uso cotidiano en el caso de Dennett parte del nivel cero de análisis: dada la concepción pragmatista, y su incorporación de la distinción creencia / opinión, el punto de partida es crítico y cuidadoso respecto de las implicaciones del lenguaje psicológico de sentido común. Para Fodor, por otro lado, el nivel de base es la aceptación de lo que nuestras expresiones de sentido común implican, exceptuando los casos raros a los que pueda darse una explicación diferente (esto es, para los que pueda proporcionarse una adecuada reformulación de la expresión problemática). Se recurre entonces a una revisión únicamente en casos que van en contra de lo comúnmente aceptado. El punto central es que, en ausencia de una explicación alternativa de este tipo, la lógica fodoriana es que *debemos* aceptar la carga contraintuitiva:

Pero si *no* estuviéramos en condiciones de proveer las traducciones adecuadas (o, aun, de visualizar cómo proveerlas), ¿qué derecho

tendríamos a considerar dichas expresiones como ontológicamente promiscuas?<sup>129</sup>.

Ahora, el problema de esto, desde la perspectiva de la distinción creencia / opinión, es que lo comúnmente aceptado es justamente el conjunto de opiniones que incorporamos más o menos conscientemente con el lenguaje de sentido común. En este sentido, el modo en que Fodor adopta el cuidado metodológico está circunscripto dentro del conjunto de preconcepciones que componen la “ciencia popular” que heredamos y, por lo tanto, no puede tener el efecto “desengañador” en virtud del que fue propuesto<sup>130</sup>. Dada la concepción intelectualista, y la no incorporación de una distinción en línea con la propuesta por Dennett, esto no puede constituir un problema: el sendero argumentativo tomado por Fodor es claramente habilitado por la concepción intelectualista del lenguaje mentalista, que, además de atribuirle un carácter homogéneo, le atribuye generalizaciones deductivas y una estructuración teórica.

La condición de Aristóteles (según la cual los objetos de las actitudes proposicionales tienen forma lógica) será ulteriormente criticada en el siguiente apartado, pero quiero hacer dos comentarios aquí. En primer lugar, quiero subrayar el hecho de que su postulación depende de que previamente se haya aceptado la primera condición. Por esta última, se aísla el objeto de las actitudes proposicionales del cual luego se especifica que tiene forma lógica. De esta manera, la crítica a la primera condición se extiende hacia lo requerido por la condición de Aristóteles. En segundo lugar, hay otro sentido en el que los requerimientos fodorianos para una teoría de las actitudes proposicionales están amparados por la concepción intelectualista del lenguaje psicológico popular. De hecho, es notable cómo la condición de Aristóteles entra en línea con la tesis de la concepción intelectualista por la que se atribuye una estructura deductiva a las generalizaciones de sentido común. Una explicación científica de la conducta por apelación a estados mentales asumirá de este modo un formato específico en tanto puede modelarse de acuerdo a las explicaciones altamente teóricas de nivel popular.

Volviendo ahora a la defensa de TAP (es decir, el análisis en términos de actitudes proposicionales tomado como estrategia general para una teoría de la mente), había antes señalado (p. 83) que una de sus ventajas más grandes es que proporciona un

---

<sup>129</sup> Fodor (1978, p. 175).

<sup>130</sup> Esto también puede relacionarse con una explicación del carácter “autoevidente” o *a priori* que Fodor atribuye a sus cinco condiciones (ver la cita en p. 86).

modelo fijo para la individuación de estados mentales<sup>131</sup> (por medio de la determinación de sus tres variables). Ahora bien, desde la perspectiva de la crítica metodológica que he expuesto, si esta ventaja se convierte en la tesis (no necesariamente explícita, pero evidentemente asumida por la teoría representacional de la mente) de que los estados intencionales son de hecho determinables fijamente, aquella se convierte en un problema metodológico grave. Esto es así en la medida en que TAP configura un objeto de estudio (esto es, las mismas actitudes proposicionales) específicamente para el cual se justifican determinados emprendimientos teóricos<sup>132</sup>. Una manera abarcadora de expresar esto es lo que Dennett (1996a, p. 56) denominó el “equivoco objetivo de la precisión proposicional”, que puede, de acuerdo con mi interpretación, entenderse como una consecuencia directa de la aceptación de TAP.

Este llamado “equivoco objetivo” se manifiesta de diferentes formas, todas relacionadas con el modo en que el lenguaje, *via* la absolución por parte de TAP, influye en nuestros intentos de configurar un modelo para la realidad psicológica que subyace al comportamiento inteligente. En general, el error para el teórico consiste en concentrarse en los productos lingüísticos para dar cuenta de los estados mentales intencionales a los que, de acuerdo con nuestras asunciones de sentido común, ellos refieren: esto es, el error consiste en concentrarse en las opiniones. Y, de acuerdo con esto, el análisis en términos de actitudes proposicionales es un modelo absolutamente apropiado, pero no para las creencias y los demás estados intencionales, sino para las opiniones.

El teórico que se concentra de este modo en los productos lingüísticos omite inevitablemente el rol activo del lenguaje en la especificación y determinación de nuestras creencias y deseos (o, más bien, de nuestras opiniones):

[N]uestro entorno lingüístico siempre nos está obligando a dar –o conceder– una expresión verbal precisa a convicciones [deseos y demás] a las que les faltan los estrechos límites con los que los dota la verbalización<sup>133</sup>.

El lenguaje, por así decir, hila demasiado fino e impone, y no refleja, una estructura y una serie de distinciones propias del mismo sistema lingüístico. Cualquier sugerencia

---

<sup>131</sup> El sentido y alcance de esta individuación tendrá una fuerte repercusión para la discusión del realismo fodoriano, que desarrollaré más adelante.

<sup>132</sup> Por ejemplo, los ya casi abandonados debates respecto del contenido estrecho o amplio de los estados mentales (que Fodor abandonó recién en *The Elm and the Expert* de 1994) y el de las atribuciones *de re* enfrentadas a las atribuciones *de dicto*.

<sup>133</sup> Dennett (1987a, pp. 31-32).

que quiera hacerse a partir de ellas para una teoría de la intencionalidad y del contenido mental estará esencialmente viciada.

Ahora bien, esta posición dennettiana no implica negar el hecho de que creencias y deseos puedan tomar una forma específica, luego de pasar por la formulación lingüística. Al contrario, la distinción entre creencias y opiniones, junto con otra distinción relacionada entre “tener creencias” (*having beliefs*) y “pensar” (*thinking*)<sup>134</sup>, justamente constituye una herramienta (poco aprovechada, en mi opinión) para *explicar* este hecho, además de proporcionar un ordenamiento del campo teórico que evita confusiones y errores heredados de nuestra costumbre de proyectar acríticamente nuestra herencia lingüística en lo que sea que se estudie. Dada la distinción, el punto de apoyo para la explicación de ese hecho es evidente: las opiniones son determinadas porque, de alguna forma u otra, *consisten* de una serie específica de palabras en nuestro lenguaje natural.

A la vez, la idea, desarrollada en el capítulo anterior, de las palabras como herramientas mentales, como entidades objetuales producto de un proceso de conformación que Dennett denomina “etiquetamiento” (*labeling*), impone un corte por el cual se entiende que el pensamiento (en tanto que actividad consciente mediada por el lenguaje, esencialmente diferente de la mera capacidad para tener creencias, deseos, etcétera) es en buena medida una actividad artificial<sup>135</sup>, no estrictamente biológica en el sentido básico en que las creencias y los deseos lo son: una actividad que, como tal, exige un aparato explicativo diferente de aquel que funcione para las creencias y los deseos. En este sentido, la distinción tiene un rol *negativo* así como también un rol *positivo* en la explicación de las atribuciones intencionales en el caso específico del ser humano (lingüísticamente competente) y de las diferencias con las atribuciones para otras especies o sistemas intencionales en general.

El rol negativo de la distinción entre creencias y opiniones tiene una importancia fundamental desde una perspectiva evolucionista. De acuerdo con el carácter secundario y derivado que, como vimos (pp. 51, 63), Dennett imputa a las opiniones respecto de las creencias, el mismo hecho de que creencias y deseos puedan asumir una forma lingüística específica es considerado un fenómeno secundario y derivado para una teoría de la intencionalidad. Si se tomara este hecho, en última instancia de orden lingüístico,

---

<sup>134</sup> Dennett (2000, p. 20).

<sup>135</sup> Bo Dahlbom (1993) es un acérrimo exponente de esta posición, que ha sido fuertemente influido por las ideas de Dennett.

como *punto de partida* de una tal teoría se estaría cometiendo un error metodológico de base, una inversión metodológica que, como trataré de mostrar, tiene claras implicaciones anti-evolucionistas: la inversión en cuestión consiste esencialmente en explicar el fenómeno básico (el que sea) concentrándose en sus manifestaciones más complejas.

Esta acusación se convierte, si se centra contra el modelo de las actitudes proposicionales, en una crítica generalizada de la pertinencia del análisis formal para una teoría de entidades, propiedades y procesos psicológicos y, *a fortiori*, en una crítica integral de TAP. Dennett expresa esta crítica muy claramente:

[L]a semántica formal nos exige determinar el objeto a evaluar en un momento especial y en un contexto especial para el valor de la verdad o la referencia, y mientras que la conducta lingüística abierta le proporciona al teórico objetos candidatos –elocuciones– para ese papel, efectuando movidas internas y postulando objetos o estados “mentales” análogos para esa determinación, debe ejercer violencia sobre la situación psicológica. Quienquiera que importe las categorías necesarias para una teoría semántica formal y las obligue a servir en una teoría psicológica, está destinado a crear un monstruo<sup>136</sup>.

Mi tesis fuerte es que esta crítica general tiene implicaciones evolucionistas, en un sentido importante.

Dennett (en Carruthers y Boucher, 1998, pp. 284-285) presenta un razonamiento hipotético que podría cometer un defensor de la tesis del lenguaje del pensamiento, y que considero que ilustra el núcleo de su crítica. El razonamiento consiste de seis pasos: (1) una de las cosas que los seres humanos hacemos es hablar uno al otro; (2) otra cosa que hacemos es hablar con nosotros mismos; (3) un refinamiento del punto 2 es hablar silenciosamente con uno mismo; (4) un ulterior refinamiento de 3 es abandonar sus características auditivas y fonéticas y pensar para uno mismo en meras proposiciones; (5) dado que las proposiciones son objetos abstractos, éstas necesitarían de vehículos físicos en el cerebro; además, dado que 4 mantiene la forma lógica<sup>137</sup> de las oraciones, debe haber un medio de representación, diferente de cualquier lenguaje natural, que tenga esta característica; (6) la actividad de manipular fórmulas de este medio representacional es la variedad fundamental del pensamiento. En este hipotético razonamiento se invierte el orden de prioridad establecido por la distinción creencia / opinión, en el sentido de que la variedad más básica de pensamiento deriva sus

---

<sup>136</sup> Dennett (1987a, p. 158).

<sup>137</sup> Exactamente como lo exigía la condición de Aristóteles (p. 85).

características de las propias de nuestra competencia lingüística y, por lo tanto, son estas características aquellas de las que debemos dar cuenta en una teoría psicológica.

Si, desde una perspectiva evolutiva, entendemos nuestra capacidad para manipular un lenguaje natural como un logro reciente y muy especializado, fruto de un proceso más o menos paulatino de diseño y perfeccionamiento gradual (¿y qué otra explicación podría haber?)<sup>138</sup>, es fácil ver cómo la lógica del razonamiento hipotético anterior va, en el fondo, directamente en contra de la teoría evolucionista. Hay dos sentidos relacionados en que esto es así, ambos a nivel metodológico: en primer lugar, la teoría evolucionista proporciona un modelo explicativo que refleja la idea de que, en general, todo fenómeno complejo lo es en virtud de un proceso de desarrollo y perfeccionamiento desde estadios más simples, con lo cual el fenómeno puede descomponerse en sus manifestaciones básicas para su explicación; en segundo lugar, y en relación con lo anterior, la perspectiva explicativa que se adopte no deberá impedir que el foco de atención se dirija a otras encarnaciones del mismo tipo de fenómenos que se pretendan explicar, pero a niveles de complejidad menores.

La teoría representacional de la mente viola ambos principios explicativos, mientras que la teoría de los sistemas intencionales los sigue. Esto es muy claro en especial para el segundo principio. Como señalaba en p. 19, la teoría dennettiana se caracteriza por una apertura metodológica, que podemos llamar el “carácter amplio” de su enfoque. Este consiste en dirigir la atención a todo tipo de organismos o sistemas no lingüísticos, tanto para estudiarlos directamente como para encontrar sustento para una explicación de los fenómenos en niveles mayores de sofisticación. En cuanto al primer objetivo, la teoría de los sistemas intencionales incorpora la preocupación de determinar las diferencias entre distintos tipos de mentes (p. 59), con lo que hay una asunción no homologadora de principio que justifica que el alcance de la investigación cubra casos no lingüísticos, como por ejemplo los niños y los animales no humanos: se plantea, en otras palabras, la cuestión explícita de las mentes no lingüísticas<sup>139</sup>. En cuanto al segundo objetivo, se abre el foco de estudio hacia otros organismos en tanto que herramienta útil y esclarecedora para nuestro caso, hablantes competentes, especialmente si tenemos en cuenta el carácter potencialmente engañoso que Dennett atribuye al lenguaje psicológico popular.

---

<sup>138</sup> Ver *The Language Instinct* (1994) de Steven Pinker para una influyente postura evolucionista sobre el desarrollo del lenguaje.

<sup>139</sup> Por ejemplo, el apartado “El Problema de las Mentes No Comunicativas” en Dennett (1996a, pp. 23-30).

Por otra parte, al igual que para el caso de la concepción intelectualista y su enfoque restrictivo, la teoría representacional de la mente puede caracterizarse como restrictiva, tanto en lo que respecta a su objeto de estudio como al enfoque explicativo adoptado. La inversión de prioridad que supone centrarse en las características lógico-semánticas de nuestro lenguaje implica que se establezca una suerte de hegemonía del tipo humano de pensamiento en la investigación. Esto por una razón trivial: los lenguajes articulados son conjuntos de herramientas de nuestra exclusiva propiedad. Este es un hecho trivial que, como tal, solemos ignorar:

Armados con esas herramientas mentales, tendemos a olvidar que *nuestras* maneras de pensar respecto al mundo no son las únicas y que en concreto no son prerequisites para enfrentarse con éxito al mundo<sup>140</sup>.

Es importante ver la relevancia evolutiva de esto: la diferencia *metodológica* que estoy marcando entre las posiciones de Dennett y Fodor encubre a la vez otra diferencia importante y general en la manera *sustantiva* de entender a los organismos vivos o bien como armados de diversos sistemas de control de la conducta para su supervivencia (para el caso del primero) o bien de entenderlos simplemente como poseedores o no poseedores de creencias y deseos (para el caso del segundo).

Dado esto, la teoría fodoriana sencillamente no contempla el problema de los diferentes tipos de mentes que puedan distinguirse. Esta no puede verse meramente como una decisión metodológica ingenua, sino que en última instancia es una grave falencia (propriadamente anti-evolucionista) de la teoría para ser sensible a las diferencias entre especies. Esto quedará claro con un ejemplo:

No se encuentran *organismos* que puedan pensar que la chica ama a John pero no pueden pensar que John ama a la chica. No se encuentran *organismos* que pueden inferir P de P&Q&R pero no pueden inferir P de P&Q. [...] Para los propósitos de este artículo, asumimos sin argumento que las capacidades cognitivas son generalmente sistemáticas en este sentido, ambos en humanos y en *muchos organismos infrahumanos*<sup>141</sup>.

Más allá del carácter burdo y anti-evolucionista de la expresión “muchos organismos infrahumanos”, es absolutamente evidente que la asunción “sin argumento” de los autores no sólo no hace justicia a las diferencias específicas entre dichos organismos, sino que los reúne bajo un mismo parámetro, aquel que nos proporcionan las

---

<sup>140</sup> Dennett (1996a, p. 173).

<sup>141</sup> Fodor y McLaughlin (1998, p. 109). La traducción y el subrayado son míos.

características lógicas y semánticas de nuestro lenguaje (como, en este caso, la sistematicidad). Lo que esto claramente muestra es el modo en que la teoría representacional de la mente, y con ella TAP, no logra hacer frente a las cuestiones relacionadas con la intencionalidad en los casos no humanos, no lingüísticos<sup>142</sup>.

El carácter restrictivo de la teoría fodoriana podría no ser tal si ésta adoptara una perspectiva diferente en cuanto al alcance de sus tesis; una perspectiva como, por ejemplo, la que adopta Donald Davidson (1975, 1982), quien directamente niega que los seres no lingüísticos sean capaces de tener estados intencionales. A pesar de la señalada falta de desarrollo del tema puntual del contenido intencional en el caso de (por lo menos algunos) animales no humanos, Fodor en algunos puntos parece dar por sentada la falsedad de la tesis davidsoniana (como, en cierta medida, en la misma cita arriba). Esto aparece muy claramente en el contexto de la refutación fodoriana de la idea de que el lenguaje del pensamiento pudiera ser simplemente un lenguaje natural cualquiera, ya que en ese caso, de acuerdo con la teoría representacional de la mente, ningún organismo no lingüístico podría tener estados intencionales<sup>143</sup>.

Ahora, en principio, pareciera que TAP favorece la tesis davidsoniana debido, por un lado, al carácter fuertemente lingüístico del análisis y, por otro lado, al hecho de que éste proporcione un modelo fijo para los estados intencionales, entendidos de este modo como un conjunto homogéneo de fenómenos. Por otra parte, no me parece excesivo decir que las intuiciones que hacen apremiante el esquema de la teoría representacional de la mente, construido como está sobre la aceptación de TAP, se derivan de modo directo de la flexibilidad lógico-semántica que nuestros lenguajes habilitan. En otras palabras, lo que hay de atractivo en la teoría radica esencialmente en cómo funciona para los casos típicos de opiniones: según lo que vengo sosteniendo, de hecho, Fodor traslada el esquema explicativo que funciona para nuestras opiniones (aunque tampoco bajo las presuposiciones idealizadas de Fodor) y admite su extensión, sin variantes ni consideraciones especiales, a los organismos no lingüísticos.

Por todo lo anterior, puede afirmarse que el llamado de atención metodológico que extrajimos de la concepción pragmatista dennettiana es de muy largo alcance. En un sentido importante, y más allá de las características particulares del aspecto lingüístico

---

<sup>142</sup> Esta es la contraparte anti-evolucionista de la apertura metodológica de la teoría dennettiana, destacada en p. 19.

<sup>143</sup> Dicha refutación se encuentra en Fodor (1975, p. 74) y en Fodor (1978, p. 190). Otro lugar en que aparece el mismo tema es en el prefacio de *Psychosemantics* (1987), donde se presenta la teoría que va a ser defendida en su aplicación al comportamiento de Greycat (el gato de Fodor, del que volveré a hablar...).

de la psicología de sentido común, lo que la distinción técnica entre creencias y opiniones brinda es, entre otras cosas, una manera de reconocer (y volver filosóficamente relevante para la teoría de la mente) el hecho obvio de que somos criaturas parlantes. El que se reconozca este hecho y sus implicaciones para una teoría de la mente es el germen de la gran mayoría de las disputas entre nuestros autores. Mientras Dennett realiza un *análisis crítico* del lenguaje mentalista que revela su función y la naturaleza de su uso, Fodor *asume* el lenguaje (y omite el análisis) y lo toma como base para la construcción de su teoría. En el siguiente apartado, estudiaré algunas de las diferencias teóricas sustantivas más relevantes que esta situación genera.

### ***3. La Primacía de la Conducta y el Realismo Intencional***

Voy a estructurar este apartado de acuerdo a dos ejes: siguiendo la distinción del realismo fodoriano hecha en el capítulo 2 (p. 36), me dedicaré primero a la cuestión del estatus de los estados mentales tipo y luego a las cuestiones referentes a su contenido intencional (lo que abarca las disputas respecto de la intencionalidad intrínseca y derivada, de la determinación o indeterminación del contenido mental y del atomismo *versus* el holismo del contenido); en ambos puntos, extraeré las implicaciones evolucionistas y anti-evolucionistas de los temas desarrollados.

Creo que es justificado abordar de esta forma concentrada las posiciones teóricas en disputa entre la teoría de los sistemas intencionales y la teoría representacional de la mente porque de hecho una de las principales tesis generales que defiendo en este trabajo es la de que existe, para cada una de dichas teorías, un conjunto compacto de presupuestos teóricos y metodológicos enfrentados que he llamado, respectivamente, enfoque evolucionista y anti-evolucionista acerca del contenido mental. Apelando a estos enfoques (y a la distinción creencia / opinión, el eje a través del cual se muestra el alcance de estos enfoques), puede atribuirse una cohesión fuerte al conjunto de tesis que conforman las teorías en cuestión y, del mismo modo, también puede darse cuenta de la incompatibilidad “en bloque” que hay entre ellas<sup>144</sup>.

Empiezo por aclarar lo que quiero dar a entender con la expresión en el título del apartado: “primacía de la conducta”. Esta es la idea básica, presente en todos los puntos

---

<sup>144</sup> Ver p. 2.

del planteo dennettiano, de que la perspectiva central desde la que hay que abordar las cuestiones acerca de la intencionalidad y los estados mentales intencionales es la del comportamiento efectivo de los sistemas objeto de atribución y todo lo inmediatamente relacionado con él. Por más vaga que esta idea pueda parecer, su importancia para la teoría de los sistemas intencionales en general no puede ignorarse. Quiero mostrar en lo siguiente cómo esta idea tiene repercusión para el problema de la configuración del concepto de “estado mental” y, a la vez, cómo la misma está implicada en el concepto técnico de creencia.

Como vimos en el anterior apartado, la aceptación de TAP puede conducir al teórico a asumir distinciones y caracterizaciones que provienen en el fondo del modo en que comúnmente hablamos acerca de nuestras mentes. Esta situación tiene una repercusión directa en el problema de definir, dentro de la teoría de la intencionalidad que se esté elaborando, lo que va a entenderse por un estado mental o intencional. Un punto central para dicho problema es el hecho (al que ya apelé en varios lugares) de que el análisis basado en las actitudes proposicionales proporciona un modelo fijo para la individuación de los estados mentales. Esto, en cierta medida, constituye una clara ventaja para el teórico. Sin embargo, si, tal como lo hace Fodor, no adoptamos el llamado de atención metodológico que extrajimos de la concepción pragmatista del discurso psicológico popular (pp. 81-82) y más bien aceptamos acríticamente el esquema (las distinciones y caracterizaciones) de este discurso, corremos el peligro de estructurar nuestra teoría, *via* la aceptación de TAP, en base a una mala concepción de lo que es un estado mental.

Para el problema específico del realismo acerca de los estados intencionales, el carácter engañoso del lenguaje de sentido común cobra un matiz “ontologizante”, por así decir. El discurso mentalista tiende a postular estados y procesos mentales individuales, y nosotros, los usuarios de ese discurso, tendemos a suponer la existencia de entidades concretas correspondientes. Esta tendencia es plenamente análoga a la advertencia wittgensteineana más general de acuerdo con la cual somos engañados por la gramática de nuestro lenguaje ordinario<sup>145</sup>. Estoy pensando en los reparos de Wittgenstein (1968, p. 27) respecto de la interpretación de la palabra “significado” en los casos en que, al exigírsenos una respuesta a la pregunta aparentemente bien planteada “¿Qué es el significado?”, irreflexivamente pasamos a buscar alguna entidad

---

<sup>145</sup> El carácter general de esta tendencia hace que su efecto exceda los casos de fenómenos intencionales como las creencias y los deseos y se extienda a casos como los dolores y el yo, por ejemplo.

para dar una respuesta al interrogante. La existencia de dicha entidad es dada por supuesta o incluso considerada necesaria en tanto que entendemos la pregunta que se nos ha hecho y en principio no tenemos razón alguna para creer que sea una mala pregunta. Esto, pensaba Wittgenstein, se debe al modo en que la gramática de nuestro lenguaje nos lleva a admitir, para algunos contextos lingüísticos, planteos que trasladamos de otros contextos.

Además de esta tendencia general, aplicada al caso del discurso mentalista, sabemos que, de acuerdo a la concepción pragmatista, nosotros, los manipuladores de opiniones, somos en una gran medida optimizadores, dotadores de sentido de nuestra conducta y de la de los demás. En este sentido, la postulación de estados y procesos mentales es, entre otras cosas, una herramienta para este emprendimiento optimizador. Esto toma la forma de una *atomización* de los procesos psicológicos: las creencias, deseos y demás estados intencionales entendidos como fenómenos discretos, separados, indivisibles. Ahora, mientras Dennett entiende esto en buena medida como una ilusión generada por el simple hecho de que hablamos sobre nuestros estados mentales, Fodor lo entiende como una manera correcta (científicamente comprobable) de hablar sobre nuestra realidad psicológica. Es importante entender que esta interpretación es la mayor fuente de apoyo para el realismo fodoriano acerca de los estados intencionales.

Si se considera el asunto intuitivamente se verá que la posición de Fodor es la más fácil de aceptar. Pre-teóricamente, podría razonarse de este modo: si es verdad que mi comportamiento está regido por lo que creo y deseo, entonces es verdad que cotidianamente tengo creencias y deseos; ahora bien, una creencia o un deseo deben de ser, en algún sentido, concretos ya que no podría tenerlos si no fueran una creencia y un deseo en particular; por lo tanto, creencias y deseos deben tener una existencia concreta que los haga ser una creencia y un deseo determinados. El que este razonamiento parezca casi obviamente correcto se debe a que se apoya en las intuiciones que heredamos de nuestra manera cotidiana de hablar sobre la mente. Lo que la teoría fodoriana proporciona, desde el punto de partida en que se asumen dichas intuiciones, es un modo de hacerlas científicamente aceptables: esto es, la tesis de que existe una (y sólo una) relación tal entre un organismo y una representación interna con contenido intencional que estar en ella es, para ese organismo, tener una determinada actitud proposicional, del tipo de las que atribuimos todos los días.

De este modo, la teoría representacional de la mente atribuye un carácter rígido al significado de los términos comunes de estados intencionales, tales como “creencia”,

“deseo”, “intención”, “duda” y demás. Aunque la teoría no necesita defender la idea de que el lenguaje psicológico popular es de hecho absolutamente coherente y consistente en el uso de dichos términos, la idea es que los términos de estados intencionales que lo componen refieren de modo por lo menos aproximado a los fenómenos discretos y salientes de los que una psicología cognitiva acabada deberá dar cuenta.

Otra manera de entender el carácter rígido que de esta forma adquieren los estados intencionales es interpretándolos en tanto que términos de clase natural, del mismo modo en que Putnam (1975) entendía términos como “agua”, “oro” y “tigre”: esto es, términos que refieren de manera unívoca a entidades particulares de nuestra ontología. Esto, como se verá más adelante, puede interpretarse como una variante de esencialismo para los términos psicológicos.

La teoría de los sistemas intencionales va plenamente en contra de una concepción de los fenómenos intencionales de este tipo y de la semántica de su vocabulario. Uno de los rasgos centrales de la concepción alternativa dennettiana es que la creencia, el deseo y los demás fenómenos intencionales que, de modo pre-teórico, consideraríamos aproximadamente fijos y unitarios, deben entenderse no como sucesos o eventos acontecidos sino propiamente como *estados* mentales intencionales. La diferencia importante que se quiere marcar es la que hay entre un acontecimiento particular, determinado, que se instancia en un cierto momento y en un cierto lugar, y un estado más o menos difuso, que no tiene necesariamente comienzo y fin, que no puede estrictamente fecharse o ubicarse espacialmente. Otra diferencia crucial relacionada es que un estado es el tipo de cosas que típicamente se pueden atribuir de modo aproximado mientras que un evento es el tipo de cosas que pueden atribuirse y confirmarse con un grado mayor de precisión. (Apelar a disposiciones puede evitar estos compromisos en mayor medida aún).

Una perspectiva alternativa para abordar el problema en cuestión es adoptada por Dennett cuando distingue los estados mentales entendidos o bien como *illata* o bien como *abstracta*<sup>146</sup>. Esta distinción ayuda a definir la clase de referente para los términos teóricos postulados: lo que esta manera de hablar agrega a la distinción entre estado y evento es la diferencia que hay, desde el punto de vista del teórico, entre postular algo por su utilidad conceptual y hacerlo con fines descriptivos. Los *abstracta* son construcciones lógicas y, en cuanto tales, su estatus ontológico depende de una

---

<sup>146</sup> Esta distinción es tomada de Reichenbach (1938).

convención (Dennett da como ejemplo los centros de gravedad, los vectores, el Ecuador). Por otra parte, los *illata* son entidades teóricas cuyo estatus ontológico debe ser definido por la teoría. Fodor claramente trata las creencias y los deseos de la psicología popular de este modo, y su definición los interpreta como discretos, indivisibles, con poder causal y susceptibles de interpretación semántica.

Respecto del estatus ontológico de los fenómenos intencionales, el foco teórico principal que separa las tesis dennettiana y fodoriana está en la interpretación que se defiende del discurso psicológico de sentido común y en la medida en que se acepten como verdaderas sus implicaciones. Con palabras de Dennett:

Es tentador suponer que cuando nos apartamos del abismo de la irracionalidad [...] la arena a la cual llegamos es la arena de la psicología popular de los pensamientos, las conclusiones, los olvidos, etc..., no meros *estados* mentales abstractos como la creencia, sino episodios, actividades o procesos concretos y mensurables en el tiempo que pueden ser modelados por constructores psicológicos de modelos y medidos y probados directamente en los experimentos<sup>147</sup>.

De acuerdo con esto, la cercanía de la tesis fodoriana respecto de nuestra visión cotidiana de los hechos mentales se corresponde con la lejanía de la tesis dennettiana respecto de esa visión, y explica el que esta última sea menos atractiva a nivel intuitivo.

Si, como sostuve en el apartado 4.1, incorporamos la distinción entre creencia y opinión y, de esta manera, individuamos en la psicología popular un nivel descriptivo máximamente optimizador, nuestra tendencia intuitiva a aceptar las categorías del discurso mentalista, junto con las connotaciones que con él estamos acostumbrados a asumir, puede identificarse como tal y considerarse teóricamente viciada. Dado esto, Dennett puede formular su crítica dirigida a la atomización teórica que antes atribuíamos a la teoría fodoriana:

[N]uestra capacidad para seguir el rastro de las cosas a través del tiempo no está bien descrita por ninguna teoría que atomice los procesos psicológicos en momentos sucesivos con ciertas características<sup>148</sup>.

Desde la perspectiva del par *illata* / *abstracta*, Dennett se refiere al uso descuidado que atribuimos a las opiniones:

La noción *corriente* de creencia coloca sin duda a las creencias en alguna parte de la mitad del camino entre ser *illata* o *abstracta*. Esto

---

<sup>147</sup> Dennett (1987a, p. 89).

<sup>148</sup> Dennett (1987a, p. 158).

me sugiere que el concepto de creencia que se encuentra en la comprensión común, es decir en la psicología popular, es poco atractivo como concepto científico<sup>149</sup>.

Ahora bien, además de esta crítica “lingüística” aplicada al problema del estatus ontológico de los fenómenos mentales, la distinción entre creencias y opiniones ofrece las bases para otro tipo de crítica a la postura fodoriana: lo que podría denominarse el “argumento de la multiplicidad de las creencias”.

La primera expresión de esta crítica de Dennett está dirigida justamente contra Fodor en la reseña de *The Language of Thought* titulada “A Cure for the Common Code?” (1977). La idea básica es que hay una diferencia entre las creencias y deseos que tenemos y las creencias y deseos que albergamos [*entertain*]<sup>150</sup>. Este sentido más fuerte de la posesión de una creencia implica su encarnación en algún medio lingüístico y, *a fortiori*, algún grado de apreciación conciente de esa creencia<sup>151</sup>. La diferencia que se quiere marcar es en el fondo el núcleo de la distinción entre creencia y opinión: las creencias que albergamos se corresponderían con las opiniones dennettianas ya que, para poder “tenerlas”, necesitamos “frasearlas” de algún modo a través del recurso lingüístico.

Dada la distinción, el argumento de la multiplicidad de las creencias consiste en que el número de las creencias que tenemos es inmensamente mayor al de las que albergamos:

Quizás “alberguemos” [*entertain*] actitudes proposicionales o bien en serie o por lo menos en números manejablemente reducidos en algún momento determinado, pero las actitudes proposicionales que *tenemos* sobrepasan por lejos aquellas que (en algún sentido) albergamos activamente<sup>152</sup>.

De este modo, existe (por lo menos) un sentido en que puede decirse de algún organismo que tiene una creencia, sin que con esto se esté implicando la instanciación de alguna relación particular con una representación mental o, más bien, con cualquier otro hecho psíquico entendible como un evento acontecido en algún contexto espacio-

---

<sup>149</sup> Dennett (1987a, p. 60).

<sup>150</sup> Esta traducción del término “*entertain*” no es quizás del todo adecuada, pero la idea que se quiere transmitir es un modo más concreto y activo de “tener” una creencia o un deseo.

<sup>151</sup> Hay que recordar aquí la estrecha conexión que Dennett traza entre la competencia lingüística y el sentido pleno de “ser conciente” para un organismo (ver nota 38 del capítulo 3 y también Dennett (1998a, pp. 358-359)).

<sup>152</sup> Dennett (1978a, p. 104). La traducción es mía.

temporalmente determinable. Un paso ulterior del argumento sería argüir que este segundo sentido de la posesión de una creencia es el fundamental.

El argumento presentado es en realidad propuesto en el marco de la crítica a la noción de representación mental de la que Fodor hace uso, y específicamente para desacreditar la idea de que nada puede ser creído o aprendido sin que sea también explícitamente representado<sup>153</sup> en algún medio. Aún así, creo que el argumento puede funcionar a los fines de, por así decir, “desontologizar” el vocabulario mentalista, en el sentido explicitado. A la vez, y sin mencionar el problema del contenido intencional, el argumento de la multiplicidad de las creencias ya da pie para defender una interpretación holista acerca de los estados mentales: esto es, una interpretación que se hace eco de la famosa tesis davidsoniana según la cual no puede atribuirse una creencia individual sin atribuir una cantidad de otras creencias relacionadas y que, por lo tanto, grandes familias de creencias viajan juntas en un mismo momento (aunque estas ideas sean defendidas en un contexto teórico diferente).

Todo lo anterior viene a coincidir con otro aspecto de la noción técnica de creencia, que destacaba en p. 63: esto es, la idea de la creencia como sustrato persistente. Esta idea subraya la importancia de distinguir entre dos estratos diferentes de información, la información perceptual y la lingüística, de los cuales el primero es el mayormente implicado en la atribución holista y no ontológicamente cargada de creencias. La información en cuestión es aquella en virtud de la cual el organismo que se esté prediciendo regula su conducta de un modo directo, no lingüísticamente mediado. No necesitamos suponer una apreciación conciente de dicha información para la atribución de estados intencionales, sino que lo crucial (lo único relevante para el intérprete) son los efectos conductuales que aquella, en algún sentido, genera. Esta es una expresión clara de lo que llamo “primacía de la conducta” y se muestra aquí en el modo en que la atribución de estados intencionales no comprometida ontológicamente es fruto de otorgar primacía a los factores comportamentales y ambientales para dar cuenta de los cuales no se necesita apelar a un formato de información diferente o más sofisticado que el meramente perceptual.

Una de las principales críticas fodorianas a esta manera de interpretar los fenómenos intencionales apunta al problema de su rol causal. Una postura que no

---

<sup>153</sup> Dado mi enfoque en este trabajo, no trato puntualmente el asunto de la representación explícita (ni de la discusión dennettiana (especialmente, Dennett, 1987a, pp. 194 y siguientes) de las representaciones implícita y tácita).

atribuya a los estados mentales una carga ontológica claramente definida, entendiéndolos como estados difusos, no discretos, no individuables, no puede, según Fodor, dar cuenta del poder causal que debemos suponer que tienen, si no queremos ver a la conducta inteligente como esencialmente azarosa y casual. Son muchos los lugares donde Fodor expone este tipo de crítica, pero me voy a concentrar en un argumento de carácter explícitamente anti-evolucionista, ya esbozado en p. 42.

La crítica al llamado “argumento evolucionista de Dennett para el principio de verdad” es, en definitiva, una crítica al instrumentalismo o interpretacionismo de la postura dennettiana, y uno de sus problemas claves, en opinión de Fodor, es la cuestión del rol causal de los estados mentales. La crítica fodoriana se reduce en el fondo a la idea de que, dado que, para Dennett, las creencias y los deseos *no existen*, tampoco podrían jugar un rol efectivo (causal) en la historia evolutiva del organismo al que se los atribuimos:

Si no *hay* creencias y deseos, luego, *a fortiori*, no puede haber nada para lo cual las creencias y los deseos hayan sido seleccionados, y no puede haber ninguna función biológica (o de otro tipo) que las creencias y los deseos desempeñen<sup>154</sup>.

Ahora bien, el problema de la crítica de Fodor es que da por supuesto un punto crucial que es altamente cuestionable. Del mismo modo en que la noción de ser verdadero asumida por Fodor es restringida (pp. 74-75), también lo es su concepción asumida de lo que puede tener un rol causal: esto es, se da por supuesto que, para que podamos considerar que algo tiene un papel causal en un determinado proceso, debemos poder identificarlo de algún modo concreto. Hay una analogía, de la que Dennett hace uso frecuente<sup>155</sup>, que en mi opinión es óptima para contrarrestar este punto: las creencias (y los demás estados intencionales) entendidas como similares a dólares. Aquí, “dólares” no hace referencia a los billetes físicos de dólares sino a la correspondiente medida de valor económico encarnada en esos billetes. En este sentido, los dólares son abstractos, no ocupan espacio físico, no tienen estructura ni se componen de partes, no acontecen en una determinada conjunción espacio-temporal y son solamente uno entre los posibles sistemas convencionales para medir el valor económico. Con todo esto, ¿estaríamos realmente justificados en decir que los dólares no tienen de hecho (o no podrían tener) papeles causales?

---

<sup>154</sup> En Dahlbom (ed.), (1993, p. 76). La traducción es mía.

<sup>155</sup> Por ejemplo, Dennett (1987a, pp. 188-189) y Dennett (1998a, pp. 328-329).

Antes de pasar a los problemas referentes al contenido intencional, quiero extraer alguna conclusión sobre las implicaciones evolucionistas y anti-evolucionistas de las dos posiciones analizadas. El principal punto evolucionista de la interpretación dennettiana de los términos mentalistas en lo que respecta a su estatus ontológico concierne al gradualismo de los estados que aquella permite. La concepción de los estados mentales que resulta de dicha interpretación es no rígida, flexible y admite una gradación continua tanto en la gama de los sistemas intencionales como en el tipo de estado en virtud de cuya atribución aquellos pueden considerarse como tales. Estas diferentes gradaciones, estos dos tipos de gradualismo, se corresponden y son interdependientes.

Por otra parte, y por las mismas razones, la concepción defendida por Fodor es rígida tanto para los conceptos mentalistas en cuanto tales como para la variedad de organismos cuya conducta puede describirse apelando a ellos: ambos sentidos en que dicha concepción es rígida pueden entenderse, acordemente al caso dennettiano, como dos tipos de anti-gradualismo. Más aún, enfrentada como está al gradualismo evolucionista, no parece inadecuado caracterizar a la postura fodoriana de esencialista, en el sentido anti-evolucionista por excelencia.

En el prefacio de *Psychosemantics*, Fodor da por supuesto que la teoría representacional de la mente, con su postulación fuerte de creencias y deseos, se aplica a su gato, Greycat, exactamente del mismo modo en que se aplica a nosotros, seres humanos adultos. Esto no quiere decir que Fodor considere plausible la tesis (obviamente errónea) de que, entre los gatos y nosotros, no hay diferencia alguna en los específicos mecanismos físicos que implementan los estados: la idea más bien es que hay una tipología determinada para la instanciación del estado, que tiene que darse por igual en los organismos en que de hecho se instancien estados intencionales.

Vista desde el enfoque de la determinación de la gama de los “verdaderos creyentes” (esto es, desde el segundo tipo de anti-gradualismo en cuestión), esta idea puede entenderse en el sentido de que la pregunta acerca de si un determinado organismo es o no un poseedor de creencias y deseos es una pregunta disyuntiva definitiva: una pregunta, en fin, que no admite matices. En otras palabras, Greycat y los demás gatos, de acuerdo con Fodor, ingresan en el círculo cerrado de los poseedores de creencias y deseos. Bien. Ahora, alguien podría preguntar: “y que hay de las ratas? Y de las ranas? Y las langostas? Y las lombrices?”. Y así sucesivamente. Ante un planteo de este tipo, Fodor supondría que tiene que haber una respuesta determinable y definitiva

para cada uno de esos interrogantes<sup>156</sup>: para todos, habría una respuesta del tipo todo o nada que nosotros, de alguna manera, tendríamos que poder determinar de modo unívoco. Finalmente, a esta univocidad corresponde la rigidez de los conceptos con los que caracterizamos los fenómenos intencionales tipo de los verdaderos creyentes, y el esencialismo acerca de ellos, como su correlato.

De modo absolutamente paralelo, la postura dennettiana admite una aplicación diferenciada de los términos mentalistas a organismos diferentes. La gradación resultante depende finalmente de lo que el intérprete puede justificar de acuerdo a sus fines y la complejidad de la conducta observada; en este sentido, dicha gradación es del todo abierta: “[l]as razones para considerar una neurona individual (o un termostato) como un sistema intencional son poco convincentes, pero no nulas...”<sup>157</sup>. Para otro enfoque de este mismo punto puede traerse a colación el homuncularismo dennettiano (pp. 19-20), de acuerdo con el cual un sistema intencional puede descomponerse en equipos jerárquicamente estructurados de sistemas intencionales (homúnculos) crecientemente menos inteligentes, en la medida que la descomposición avanza. Y, de nuevo, así como la noción de sistema intencional es abierta y difusa, también lo es la noción de estado intencional que le subyace.

A la vez, esta noción difusa, característicamente evolucionista, de estado intencional, logra ser sensible a los diferentes grados de inmersión del agente en el mundo, inmersión que en última instancia se revela en términos de su flexibilidad conductual. En el fondo, el principio de la primacía de la conducta tiene por sí mismo un matiz evolucionista. Hay de hecho un vínculo estrecho entre la centralidad otorgada al intercambio directo agente / mundo para la atribución de estados intencionales y el carácter evolucionista de la postura de Dennett a este respecto. El punto de unión que hace corresponder el carácter gradualista de la noción de estado intencional con el carácter abierto de la noción de sistema intencional es la defensa de una perspectiva atributiva centrada en las complejidades y diferencias conductuales entre distintos organismos. Por otra parte, para el caso del anti-evolucionismo de la teoría representacional de la mente, la escasa relevancia concedida a la diversidad y complejidad de las conductas de los organismos refleja una mala aptitud para dar cuenta de las diferencias comportamentales entre diversos seres vivientes.

---

<sup>156</sup> En el artículo de 1986, “Why Paramecia Don’t Have Mental Representations” (“Por qué los Paramecios No Tienen Representaciones Mentales”), Fodor justamente se encarga de dar una de estas respuestas.

<sup>157</sup> Dennett (1998a, p. 362). La traducción es mía.

Doy aquí por concluida mi discusión del estatus ontológico de los estados mentales y paso ahora a tratar algunos puntos acerca del problema, importante y complejo, del contenido intencional de esos estados. En “Fodor’s Guide to Mental Representation” (1985), el filósofo llega a un punto, a su parecer, fundamental:

[L]o que mayormente motiva el Anti-Realismo es algo más profundo que la especulación empírica de que las explicaciones de la psicología de sentido común no se revelarán como científicas; más bien, es la impresión de que hay algo intrínsecamente erróneo respecto de lo intencional<sup>158</sup>.

Estoy en total acuerdo con Fodor en que el debate entre su posición y una posición instrumentalista como la de Dennett en gran medida gira alrededor de un desacuerdo sustancial respecto de la concepción de la intencionalidad. En particular, hay una conexión fuerte entre los problemas que el teórico del contenido mental entienda como justificables y significativos, por una parte, y la particular concepción de la intencionalidad que se defienda (o que se dé por supuesta), por otra. La discusión respecto de dicha concepción nos lleva nuevamente a la disputa entre la defensa de una intencionalidad intrínseca u original y su negación.

Ya he presentado la disputa en cuestión en el capítulo 1 (pp. 23-24), exponiendo el argumento evolucionista dennettiano en contra de la tesis de la intencionalidad intrínseca, y en el capítulo 2 (pp. 35-36), interpretando T2, con su postulación de representaciones mentales con contenido, como una defensa de dicha tesis. Aquí quisiera nada más destacar el modo en que el evolucionismo puede tener una incumbencia directa e importante en esta disputa. Por la vía positiva, creo que es posible interpretar la tesis de la intencionalidad derivada, en tanto que único tipo de intencionalidad posible, como intrínsecamente evolucionista, en el sentido fuerte de que es *necesario* apelar a consideraciones evolucionistas para su defensa plena; por la vía negativa, y por las mismas razones, creo que es posible interpretar la tesis de la intencionalidad original como intrínsecamente anti-evolucionista, en el sentido fuerte de que va directamente en contra del planteo teórico evolucionista. Quisiera brevemente mostrar por qué creo que ambas interpretaciones no son sólo posibles, sino correctas.

Creo que la asunción argumentativa que todos los defensores de una intencionalidad intrínseca u original comparten es entender que, para la explicación de la intencionalidad, no hay una posibilidad teórica intermedia entre la opción de un

---

<sup>158</sup> Fodor (1985, p. 83). La traducción es mía.

regreso infinito y la de un fundamento intrínseco, esto es, alguna entidad intencional en primera instancia, algo similar, en fin, a un motor inmóvil aristotélico. Este es un argumento erróneo que se repite constantemente en la historia de la filosofía: sencillamente, no se concebían posibilidades alternativas. De acuerdo con él, se razonaría de esta forma: a menos que haya alguna intencionalidad intrínseca, no derivada, no podría haber intencionalidad derivada alguna ya que ésta, por definición, se deriva de otra intencionalidad, que se deriva de otra y así sucesivamente: el regreso tiene que parar en algún punto no derivado de ningún otro.

Ahora, la importancia de la teoría evolucionista, y lo absolutamente novedoso que hay en ella a nivel del esquema explicativo que propone, es justamente que da sentido a una tercera posibilidad teórica para explicar el origen de muchas de las características más complejas de este mundo: esto es, un *regreso finito*, donde la característica que se intenta explicar se descompone en un proceso gradual que, si es invertido, va desde una instancia nula a instancias siempre más sofisticadas de la misma.

La negación de esta posibilidad explicativa es, sin más, una forma de esencialismo. Considérese este argumento analógico dennettiano:

[C]ada mamífero tiene a un mamífero como madre –pero esto implica una genealogía infinita de mamíferos, lo cual no puede darse. La solución no es buscar una *esencia* de lo mamífero que nos permitiría en principio identificar el Primer Mamífero, sino más bien tolerar un regreso finito que conecte los mamíferos a sus ancestros no mamíferos a través de una secuencia que sólo puede ser fragmentada arbitrariamente<sup>159</sup>.

Lo que se sigue del argumento es obvio para el caso en cuestión: la tesis de la intencionalidad original, en tanto niega la posibilidad explicativa evolucionista, es, sin más, una expresión contemporánea de esencialismo. Y la noción de representación mental que Fodor invoca para dar cuenta de la intencionalidad, tal como él la entiende, es, por lo mismo, intrínsecamente esencialista.

Ahora, si, como hice en la presentación de T2, conectamos la defensa fodoriana de la intencionalidad original con su realismo intencional y, específicamente, con su postulación de entidades discretas portadoras de contenido intencional (es decir, el representacionalismo fodoriano), estaríamos trazando un vínculo importante entre su enfoque anti-evolucionista y su perspectiva realista acerca del contenido mental. No

---

<sup>159</sup> Dennett (1998a, p. 362). La traducción y el subrayado son míos.

sólo esto, sino que para el caso de Dennett se sostiene el razonamiento exactamente paralelo: un enfoque evolucionista acerca de la intencionalidad conlleva de modo necesario algún grado de *anti-realismo* acerca del contenido mental.

La postura dennettiana acerca del contenido mental es funcionalista. Lo que hace que algo sea intencional es, en cada caso, su función. Ahora, ¿es la función de algo determinable independientemente del contexto más amplio en que se encuentra? No. Justamente, lo que hace que ese algo sea funcional es su relación con ese contexto. Y, más aún, el contexto hacia el que me dirijo para determinar la función de algo es necesariamente un contexto histórico de uno u otro tipo: para el caso de un artefacto de manufactura humana, debo apelar a los hechos históricos sobre el proceso de su elaboración y sobre su utilización corriente; para el caso de un ser viviente, debo apelar a su historia evolutiva y a su “uso” corriente<sup>160</sup>. De uno u otro modo, la función de algo nunca es determinada intrínsecamente, sino que es determinada “desde afuera”, poniendo en relación ese algo con el contexto mayor en el que opera<sup>161</sup>. Si uno acepta esto, está automáticamente incorporando a su teoría del contenido algún grado de atribucionismo (o, finalmente, de anti-realismo).

El atribucionismo es la tesis de que la determinación del contenido de un estado mental es extrínseca, en el sentido de que no depende, ni puede depender, únicamente de características internas de los mecanismos cognitivos de un organismo. En ausencia de una determinación interna de este tipo, el atribucionismo implica aceptar, entre otras cosas, que el contenido mental es *de por sí* indeterminado. La determinación es atribuida. Y el único medio que permite una atribución precisa del contenido intencional de un estado mental es el que nos proporcionan nuestros lenguajes naturales. Es en tanto que somos manipuladores de opiniones que podemos atribuir contenidos determinados a los estados mentales.

Quiero traer nuevamente a colación aquí la idea (presentada en p. 48) del lenguaje como una suerte de modalidad perceptiva y de las palabras como objetos fijos que condensan y simplifican el conjunto complejo de características de las experiencias

---

<sup>160</sup> Vale recordar aquí que el carácter (meramente) histórico de la explicación evolucionista era uno de los motivos principales por los que Fodor la juzgara como irrelevante (ver p. 42).

<sup>161</sup> Nótese cómo, para el caso de la atribución de contenidos intencionales, esta “puesta en contexto” ya implica una atribución holista de significados, que vincule el contenido de un estado particular con (algunas de) las demás creencias y deseos del organismo interpretado. Por otro lado, de modo análogo, la posibilidad de una determinación intrínseca del contenido implica de por sí la tesis del atomismo del contenido, que Fodor (1998) enérgicamente defiende: una opción holista, de hecho, haría depender el contenido de un estado particular de sus múltiples y variables conexiones con otros elementos semánticos involucrados.

a las que nosotros (y nuestros antepasados) nos enfrentamos. Este carácter fijo de las palabras, fruto del proceso que Dennett llama “etiquetamiento”, es el responsable de que cualquier atribución lingüística de contenido intencional sea, hasta cierto punto<sup>162</sup>, determinada, unívoca. El problema está en pasar de la afirmación de que la atribución intencional es, de este modo, determinada (dadas las características del lenguaje) a la de que existe un estado interno al que refiere dicha atribución, que revela de manera unívoca el contenido intencional atribuido (u otro diferente, pero también determinado, en caso de que la atribución sea errónea).

El paso implícito de la primera a la segunda afirmación puede revelarse como injustificado apelando nuevamente a la advertencia metodológica dennettiana respecto de las implicaciones lingüísticas, presentada en el apartado 4.1. Este rol preventivo de la distinción entre creencias y opiniones se reitera entonces para el problema de la determinación del contenido, y se expresa en la idea de que, más allá de una defensa, explícita o no, de la tesis de la intencionalidad intrínseca, existe una tendencia injustificada por la que, de la rigidez y determinación propias de nuestras atribuciones intencionales lingüísticas, pasamos a concebir los estados mentales en cuanto tales, o mejor, a sus contenidos intencionales, como semánticamente determinados.

Este profundo error nos lleva a embarcarnos en emprendimientos teóricos severamente descarrilados en tanto que se asientan y dependen de él: dos casos claros de esto en la teoría representacional de la mente son el “problema” de la naturalización del contenido (no hay tal problema, si el contenido no es visto como una propiedad intrínseca de algún estado físico) y el “problema” de la disyunción (no hay tal problema, si no se considera que haya una y sólo una interpretación correcta del contenido de un estado intencional)<sup>163</sup>.

Ahora bien, el aporte más importante de la distinción entre creencias y opiniones que quiero destacar para la cuestión del contenido intencional no es, como el recién mencionado, de tipo metodológico sino plenamente sustantivo. Una de las características centrales que cobran las creencias (en el sentido técnico de la distinción) es que están en una relación directa con la conducta, la cual, desde la actitud intencional, tratamos de interpretar. Esta característica da pie para la expresión más

---

<sup>162</sup> No hay acuerdo en la literatura respecto de si la tesis de la indeterminación acerca del contenido mental es co-extensiva con la tesis quineana de la indeterminación de la traducción radical, aplicada a la semántica del lenguaje natural. Dennett mantiene que una conlleva la otra (1987a, p. 47), mientras Bennett (1993, p. 381), por ejemplo, la rechaza.

<sup>163</sup> Como vimos en el apartado 2.1 (respectivamente, pp. 36-37 y p. 39), ambos problemas son capitales de acuerdo al planteo fodoriano.

clara de lo que llamo la “primacía de la conducta” en la teoría de los sistemas intencionales.

En pp. 62-63, caracterizamos el conjunto de operaciones que realizamos sobre las opiniones como esencialmente “liberal”, principalmente a partir del hecho de que existe un hiato entre lo que puede considerarse la posesión de opiniones y las posibilidades conductuales efectivas que dichas opiniones pudieran de algún modo especificar o restringir. Por otra parte, definíamos las creencias como directamente implicadas en las alternativas conductuales adoptadas por el sistema intencional interpretado; la contraparte de esto es que la postulación de creencias dependía directamente de la observación de la conducta del organismo y de su entorno ambiental.

La cercanía entre la creencia y la conducta es un rasgo absolutamente crucial del planteo dennettiano y, más aún, de su carácter evolucionista. Por más que esta relación entre creencia y conducta pueda parecer obvia en una primera instancia (¿quién negaría, después de todo, que las creencias son ese tipo de cosas a las que apelamos para dar cuenta de la conducta inteligente de los organismos?), puede sin embargo ocupar un rol más o menos central para una teoría del contenido, y las teorías dennettiana y fodoriana son ejemplos claros del alto grado en que esta diferencia puede darse en un contexto explicativo filosófico. Mientras Fodor pretende dar cuenta del hecho de que un estado mental adquiera una interpretación semántica o, más bien, un contenido intencional con total independencia de la relación que éste tenga con la conducta del organismo que se encuentra en dicho estado, Dennett no ve otro modo de dar sentido a este mismo hecho como no sea a través de una apelación a las disposiciones conductuales de ese organismo.

Desde la perspectiva dennettiana, no sería equivocado decir que lo único que en realidad estamos implicando cuando imputamos a alguien una creencia determinada es que ese organismo está dispuesto a conducirse de cierto modo dadas ciertas condiciones (esto es, dada su “situación ambiental” (pp. 9-10)). En este sentido, lo que en último término delimita el significado que algo en el mundo pueda tener para una criatura es su repertorio conductual. La semántica de los estados intencionales de un organismo depende directamente de la variedad y complejidad de las relaciones, en última instancia conductuales, que ese organismo mantiene con su entorno. De este modo, los modelos explicativos posibles para dar cuenta de su conducta pueden volverse más y más específicos (más y más determinados a nivel semántico) en la medida en que ésta sea más y más compleja. El sentido en que esto es así es absolutamente paralelo al

sentido en que, para Dennett, las propiedades semánticas de un estado son determinadas únicamente en virtud de los papeles funcionales que éste cumple.

La funcionalidad biológica de un estado mental consiste primariamente en el rol que éste tenga en la regulación de la conducta y, a la vez, el contenido intencional de ese estado (no determinado de por sí sino atribuido desde cierto enfoque explicativo / predictivo) es definido únicamente en base a la observación de dicha conducta y a la medida en que se entienda que ésta es adecuada a la “situación ambiental” del organismo:

[S]ólo se puede dar una razón de ser a la asociación de mensajes o contenidos verbales con acontecimientos y estados señalando la contribución efectiva de estos acontecimientos y estados a la dirección de la conducta que es, en última instancia, adecuada para la supervivencia del sistema como organismo del mundo<sup>164</sup>.

La determinación del contenido de una creencia enfocada desde el punto de vista de la contribución que ésta tendría sobre la conducta es claramente afín al modelo de explicación evolucionista, centrado como está en la actuación de los organismos y el modo en que los mecanismos que la regulan les permiten desenvolverse satisfactoriamente en su entorno.

La postura fodoriana en algunos sentidos puede verse como diametralmente opuesta a este modo de enfrentar el problema del contenido mental. El contenido mental, entendido como real y determinado, es concebido en estricta independencia del comportamiento y, consiguientemente, de la función biológica. En lugar de un enfoque de este tipo, Fodor apela, como sabemos, a una noción fuerte de representación mental, una noción que debe cumplir un rol que es tanto anti-biológico como lo es anti-evolucionista. Fodor carga a las representaciones mentales con la tarea de dar cuenta de la capacidad de responder a los significados que suponemos cotidianamente en nuestro discurso mentalista (entendido en cuanto impregnado de ideología y no como habilidad). Este es, en el fondo, el núcleo de la condición de Aristóteles, presentada en pp. 85-86.

Como ya destacué (p. 89), la condición de Aristóteles depende de que ya se haya aceptado que hay objetos mentales intrínsecamente portadores de contenido intencional (esto es, las representaciones mentales fodorianas). Ahora, dada esta concepción de los contenidos mentales como entidades objetuales, Fodor puede pasar a postular relaciones

---

<sup>164</sup> Dennett (1969, p. 123).

causales entre ellos, esto es, entre las mismas representaciones, relaciones en virtud de las cuales la forma lógica pueda mantenerse (como lo especifica la condición) y, de modo derivado, pueda finalmente legitimarse el conjunto de atribuciones intencionales que conforman nuestra psicología de sentido común. De acuerdo con esta postura fodoriana, la semántica de los estados intencionales de un organismo da cuenta del sentido mínimo y obvio en el que entendemos que las creencias y los deseos de la psicología popular regulan nuestras acciones (esto es, nuestra conducta) de un modo extremadamente directo. Según esto, las operaciones en último término semánticas entre estados intencionales *producen* la conducta en un sentido muy real, no metafórico.

El sentido directo en que, de acuerdo a la condición de Aristóteles y en virtud de la noción fodoriana de representación mental, las relaciones semánticas de los procesos mentales se corresponden con sus relaciones causales (esto es, los procesos computacionales de T3 (p. 37)) da una imagen extremadamente idealizada del funcionamiento de la mente y, sobre todo, una imagen de una nitidez y transparencia lógicas inaceptables desde un punto de vista biológico. Ya sabemos, por la discusión en el apartado 4.1, que esto se debe en gran parte a los requisitos que el teórico que toma las explicaciones psicológicas populares como un relato fiel de procesos psicológicos internos se ve compelido a satisfacer.

Ahora, este “logicismo” supuesto en las operaciones de la mente solamente puede sostenerse dentro de un enfoque anti-evolucionista del contenido mental, y es por este motivo que sus implicaciones más fuertes pueden ser entera y explícitamente incorporadas a la teoría representacional de la mente:

Mantener que los procesos mentales de un organismo son un modelo de una lógica [...] no equivale entonces a mantener que hay una creencia del organismo que corresponda a cada teorema de la lógica [...]. Equivale *solamente* a mantener que los postulados de la lógica son *representados mentalmente* por el organismo, y que esta representación mental contribuye (de modos apropiados) a la *causación* de sus creencias<sup>165</sup>.

Frente a lo que llamé la “primacía de la conducta” dennettiana, llamaría a este tipo de enfoque “*primacía de la conciencia*”. Aunque Fodor nunca apele en su teoría a las cuestiones relativas a la conciencia ni nunca se ocupe estrictamente del problema de la conciencia<sup>166</sup>, sospecho que hay un conjunto tácito de intuiciones vinculadas al modo

---

<sup>165</sup> Fodor (1981, p. 120). La traducción y el subrayado son míos.

<sup>166</sup> Una actitud que, por otra parte, el mismo Dennett (1987a, p. 12) critica.

en que pre-teóricamente la conciencia es entendida. Esto se refleja en parte en la “autoridad” que tradicionalmente se supone que tenemos sobre nuestros estados intencionales o, para el caso fodoriano, en el modo directo en que se entiende que las operaciones computacionales, que finalmente manipulan propiedades semánticas, dan lugar al comportamiento, entendido así como efecto o producto directo de causas mentales. Pero, fundamentalmente, dicha primacía se refleja en el hecho de que el único tipo de información relevante para dar cuenta de las operaciones de la mente sea modelado sobre el *relato*, esto es, sobre el tipo de conocimiento proposicional (lingüístico) del que típicamente somos concientes<sup>167</sup>.

A pesar de que, entonces, Fodor ciertamente establece una relación entre estados intencionales y conducta, su explicación del proceso de determinación del contenido de esos estados prescinde totalmente de un recurso a esta última: la apelación a la noción fuerte de representación mental justifica el que este recurso sea del todo prescindible. Ahora bien, esta divergencia en el foco explicativo genera una clase de explicaciones de los procesos cognitivos que es severamente desencaminada en la medida en que es idealizada y anti-biológica. Para concluir, quiero ilustrar que esto es así utilizando uno de los tres casos de teorías psicológicas cognitivas, presentadas por Fodor en el primer capítulo de *The Language of Thought*<sup>168</sup>, que mostrarían la necesidad de apelar a una noción fuerte de representación mental: elijo el caso de las teorías de la decisión o de la acción deliberada (y no las demás: las teorías del aprendizaje de conceptos y de la percepción) por la evidente cercanía que éstas suponen entre la conducta y los procesos psicológicos.

El modelo, presuntamente consensuado por las teorías psicológicas de la cognición, de la explicación de la acción deliberada es el siguiente: (1) el agente se encuentra en la situación S; (2) cree que en S hay un conjunto de opciones conductuales disponibles; (3) establece un conjunto de hipótesis sobre las consecuencias que se seguirían de seguir cada opción; (4) asigna un orden de preferencia a estas consecuencias; (5) elige y realiza una conducta en función de estas preferencias. El supuesto del que Fodor parte para presentar este modelo es de por sí altamente cuestionable:

---

<sup>167</sup> Cabe destacar nuevamente el rol del concepto de “opinión” en cuanto que identifica este tipo de conocimiento y también el de la distinción creencia / opinión en cuanto que subraya la prioridad teórica de las cuestiones acerca del contenido por sobre las de la conciencia (ver p. 64).

<sup>168</sup> Fodor (1975, pp. 48-69).

Doy por supuesto que es evidente en sí mismo [¡!] que los organismos creen muchas veces que la conducta que producen es conducta de una clase determinada y que ello constituye con frecuencia parte de la explicación de la forma en que se comporta un organismo para atenerse a las creencias que tiene sobre la forma de la conducta que produce<sup>169</sup>.

El problema de esto es que, más allá de lo poco “evidente en sí mismo” que lo que se supone pueda parecer, ya está implicando lo que se quiere probar. Incluso dejando de lado esto, el modelo presentado para describir los procesos computacionales es, de modo muy claro, una extrapolación de los procesos de decisión conciente característicos de un hablante competente<sup>170</sup>.

Este tipo de explicación es un ejemplo evidente de lo que denominaba “primacía de la conciencia”, que no es una perspectiva asumida por Fodor pero, como argumenté, puede atribuírsele. Ahora, en la medida que el *tipo* de explicación requerida por Fodor es viciada, modelada como está sobre el relato explícito de información, rechazo la idea fodoriana de que la adecuación de la teoría representacional de la mente depende a su vez, en una primera instancia, de la adecuación de las teorías psicológicas empíricas sobre las que se modela: fundamentalmente, me remito al argumento general (desarrollado en los previos apartados) por el cual la adecuación de dicha teoría, así como de la teoría dennettiana, puede anclarse en una determinada interpretación del discurso de la psicología de sentido común, y *no* en la psicología científica. Ahora, la cuestión puntual de la relación que las teorías dennettiana y fodoriana mantienen con otras disciplinas será tratada en lo siguiente, siempre desde el punto de vista de su respectiva inserción en los enfoques evolucionista y anti-evolucionista acerca del contenido mental.

#### ***4. Implicaciones Meta-Teóricas y la Relación entre Niveles Explicativos***

En este último apartado, quisiera hacer algunas breves observaciones sobre el aspecto meta-teórico de la postura dennettiana, como siempre contrastándola a este respecto con la teoría fodoriana: esto es, me referiré a algunos de los diferentes tipos de

---

<sup>169</sup> Fodor (1975, p. 98).

<sup>170</sup> Téngase en cuenta también que el planteo de Fodor es pretendidamente aplicable tanto a la psicología de los seres humanos como a la de otros organismos inteligentes (¡aunque, como sabemos, no importe demasiado determinar *cuáles!*). Fodor acepta esto explícitamente (1975, p. 74).

investigaciones aplicadas que son mayormente favorecidas por las teorías y, sobre todo, al modo en que puede considerarse que éstas de algún modo configuran el formato de aquellas<sup>171</sup>. Para hacer esto, tendré en consideración también otro asunto, absolutamente vinculado con este, y que ya analicé en detalle en el apartado 4.1.

El problema del tipo de explicación de los fenómenos y procesos psicológicos cognitivos que Dennett, dada su teoría, y Fodor, dada la suya, creen posible y, consecuentemente, el problema del tipo de investigación que consideran pertinente y justificada para los fines aceptados, deben encararse en una primera instancia considerando los sentidos radicalmente diferentes en que las teorías en cuestión conciben el carácter ineludible del discurso mentalista de sentido común y su pertinencia para los emprendimientos científicos. Como mencionábamos al comienzo del capítulo, las cuestiones acerca del discurso mentalista (fundamentalmente, las que llamé “concepciones” de la psicología popular) rozan en último término todos los aspectos de las teorías de la intencionalidad aquí estudiadas.

Como primera aproximación, puede decirse que la teoría de los sistemas intencionales y la teoría representacional de la mente proyectan una visión o perspectiva meta-teórica general fuertemente divergente. Esta diferencia se corresponde con el enfoque teórico que en el previo apartado denominé “primacía de la conducta” en tanto que enfrentado al enfoque opuesto de la “primacía de la conciencia”. La repercusión que esto tiene a nivel meta-teórico se traduce en la diferencia básica que hay entre concebir la investigación empírica en tanto que regulada mayormente en términos de capacidades y concebirla en tanto que regulada mayormente en términos de la arquitectura cognitiva (innata) de la mente. Esta diferencia, difusa y aproximada como es, de por sí no determina ni define el tipo de investigación en cuestión, pero aún así traza una dirección o un marco general, por el cual se enfatiza la importancia de un conjunto de aspectos por sobre otro. En este sentido, no es una diferencia menor, y merece atención como tal.

Es importante tener en claro que la diferencia que estoy marcando no implica negar que, en última instancia, tanto Dennett como Fodor tengan en la mira fundamentalmente el mismo objetivo teórico: esto es, la construcción de modelos adecuados de los procesos psicológicos en virtud de los cuales los organismos actúan. Ahora bien, es claro también que hay diversos y enfrentados modos en que puede

---

<sup>171</sup> No es mi intención hacer un desarrollo profundo de las posiciones dennettiana y fodoriana respecto de este punto; mantendré el análisis a un nivel deliberadamente general a los fines de relacionar mis observaciones con los enfoques metodológicos evolucionista y anti-evolucionista.

concebirse la realización de esta tarea (incluso hay diversos grados en que el aporte filosófico mismo pueda considerarse relevante o pertinente a este respecto)<sup>172</sup>.

Las teorías dennettiana y fodoriana de la intencionalidad tanto *especifican* algunas vías en que esta tarea pueda realizarse como *restringen* el espectro de las posibilidades para ese objetivo. Ahora, el punto relevante es que dichas teorías hacen esto en grados y de maneras sumamente dispares. Mi idea es que también esta disparidad en cuanto al aporte meta-teórico puede explicarse en buena medida apelando a la inserción de la teoría de los sistemas intencionales en el enfoque evolucionista acerca del contenido mental y la de la teoría representacional de la mente en el enfoque anti-evolucionista.

Más allá de las características específicas del tipo de emprendimiento científico que estas teorías sustentan o apoyan (esto es, lo que más arriba distinguí a grandes rasgos como un énfasis en las capacidades de los organismos, por un lado, y en su arquitectura cognitiva, por otro), la diferencia más grande entre ambas teorías es que, mientras la propuesta dennettiana es mayormente *aproximativa* en sus implicaciones meta-teóricas (es decir, esboza un rumbo investigativo), la propuesta fodoriana es *estipulativa* (es decir, define un modelo investigativo). Esto es especialmente claro en este último caso, en tanto que toda la construcción teórica fodoriana es explícitamente modelada en la cercanía de, o incluso de forma paralela respecto de, “nuestra mejor ciencia”<sup>173</sup> o, incluso, “los únicos modelos psicológicos de los procesos cognitivos que parecen ser al menos remotamente plausibles”<sup>174, 175</sup>. Por otra parte, Dennett no intenta especialmente especificar los posibles caminos teóricos que el científico puede querer recorrer, sino que asume una perspectiva meta-teórica amplia; a la vez, hay también un importante aspecto restrictivo en su postura por el que algunos de esos caminos son plenamente descartados.

De acuerdo con esto, habría que mantener que el evolucionismo mismo, como enfoque metodológico general de la teoría de los sistemas intencionales, es mayormente aproximativo. Creo que esto es en general correcto, aunque en ningún modo debiera interpretarse en sentido peyorativo. Por una parte, la teoría dennettiana no proporciona un modelo de explicación de los fenómenos intencionales que siquiera se aproxime al

---

<sup>172</sup> En la Conclusión, hago algunas breves reflexiones sobre este punto meta-filosófico.

<sup>173</sup> Fodor (1987, p. 39).

<sup>174</sup> Fodor (1975, p. 47).

<sup>175</sup> Esta idea es virtualmente el *leitmotiv* argumentativo de *The Language of Thought*, obra concebida como una suerte de destilación de los supuestos ya asumidos por las teorías psicológicas y lingüísticas más prometedoras del momento.

grado de desarrollo y detalle del que propone la teoría de Fodor, pero, por esto mismo, aquella es comparativamente mucho más fértil en lo que respecta a las posibilidades investigativas que favorece o admite. Para decirlo en una frase, la evolución no nos da todas las respuestas, pero puede generar un marco general dentro del que muchos emprendimientos teóricos diferentes pueden coexistir y beneficiarse mutuamente.

Por otra parte, es a la vez imprescindible que este marco se genere, en tanto que también puede poner restricciones y límites importantes al teórico. El sentido en que, como mencionaba arriba, el enfoque evolucionista dennettiano es restrictivo en lo que respecta a su propuesta meta-teórica substantiva, no es menor, sino que su importancia radica en que nos cierra algunas puertas que llevan a emprendimientos teóricos inherentemente viciados, dados algunos de sus supuestos iniciales de base. Ahora, como mencioné al comienzo del apartado, puede darse cuenta del carácter meta-teórico de la teoría dennettiana apelando a su interpretación del discurso mentalista de sentido común<sup>176</sup>. De esto me ocuparé en lo que sigue.

Quiero recordar mi estrategia argumentativa en el apartado 4.1. Allí defendí una interpretación de las teorías dennettiana y fodoriana que identificaba la raíz de las disputas particulares entre ellas (y el motivo principal de su incompatibilidad general) en la divergencia entre las concepciones del lenguaje psicológico popular que aquellas, de un modo u otro, aceptan o asumen. Dada esta divergencia, la teoría de los sistemas intencionales refleja una postura meta-teórica de acuerdo con la cual la investigación psicológica, del tipo que sea, se mantendrá neutral respecto de la implementación efectiva de los estados y procesos mentales tal como son identificados, descritos y caracterizados por el discurso mentalista de sentido común.

Por otra parte, la teoría representacional de la mente interpreta las leyes de una psicología científica acabada como una continuación y reelaboración de las generalizaciones altamente teóricas de dicho discurso. De este modo, para el caso de ambas teorías de la intencionalidad, hay un vínculo estrecho entre una determinada interpretación de la psicología popular y del vocabulario intencional parte de ella, por un lado, y una determinada postura meta-teórica respecto de la relación entre diferentes niveles explicativos, el nivel personal y el sub-personal, por otro.

---

<sup>176</sup> Sin ir más lejos, en parte ya hemos hecho esto en la medida que el análisis en términos de actitudes proposicionales es interpretado, como lo hace Fodor, como un modelo explicativo para la psicología cognitiva, y en la medida que su crítica por parte de Dennett se asienta, como sostuve, en una determinada interpretación de la psicología popular.

Como ya sabemos, ambas teorías de la intencionalidad son anti-eliminativistas respecto del vocabulario mentalista. Más específicamente, un punto anti-eliminativista compartido tanto por Dennett como por Fodor es que ambos consideran científicamente ineludible (e insustituible) el tipo de explicación dada en términos intencionales. En otras palabras, ambos autores otorgan una relevancia significativa al problema del alcance que puede justificadamente cobrar la incorporación y el rol del discurso intencionalista en una psicología científica y en otras disciplinas relacionadas.

Ahora bien, como sostenía, la medida importante en que difieren las concepciones de la psicología popular supuestas por los filósofos tiene una fuerte repercusión en este problema: así, el modo en que el discurso mentalista será incorporado y jugará su rol en las diferentes disciplinas científicas diferirá considerablemente y, de acuerdo con esto, diferirá también el tipo de explicación considerada pertinente y justificada. A modo ilustrativo, consideremos el caso del carácter causal de las explicaciones formuladas en términos mentalistas.

Para Dennett, y dada la concepción pragmatista de la psicología popular, puede considerarse que las explicaciones del comportamiento que formulamos cotidianamente apelando a estados y procesos psicológicos identificados en términos mentalistas son causales solamente en un sentido débil<sup>177</sup>:

“Si Tom no hubiera creído que  $p$  y hubiera querido que  $q$ , no habría hecho A.” ¿Es ésta una explicación causal? [...] Se le puede llamar explicación causal porque habla de causas, pero es seguramente todo lo poco específica e inútil que una explicación causal puede llegar a ser<sup>178</sup>.

En la medida en que esto sea así, no estamos forzados a dar explicaciones ulteriores de este sentido debilitado de “causal”, en el contexto en que se da.

Por otra parte, para el caso de Fodor, el carácter causal de las explicaciones de sentido común es considerado una de sus características esenciales y, más aún, una condición necesaria para que un fragmento de discurso pueda interpretarse correctamente como un ejemplo de psicología popular (ver p. 31). La idea de “causal” aquí (así como la de “causación mental” entre muestras de representaciones mentales) es en ningún modo deflacionaria, sino plenamente “específica” y “útil”, y, además, de una importancia radical: es la idea que está detrás del paralelismo, que, según Fodor, el discurso intencionalista supone, entre las relaciones causales entre estados mentales y

---

<sup>177</sup> Ver p. 103.

<sup>178</sup> Dennett (1987a, p. 62).

las relaciones semánticas entre sus objetos proposicionales. La misma teoría representacional de la mente es entendida por su autor como una manera de dar cuenta de este fundamental paralelismo (ver p. 33).

Una de las ideas básicas más profundamente arraigadas en el pensamiento dennettiano (ya fuertemente presente en *Content and Consciousness* de 1969) es la de que el nivel explicativo desplegado en el vocabulario intencional debe interpretarse en términos de “interpretación ulterior”<sup>179</sup> respecto de las descripciones físicas y fisiológicas que puedan hacerse a nivel sub-personal. Esta es una sofisticación de la radical idea quineana de que el empleo del vocabulario intencional debe verse como un “modismo esencialmente dramático” (*a dramatic idiom*)<sup>180</sup>. Ahora bien, la explicación mentalista puede aportar un nivel de descripción diferente, que viaja paralelamente, pero no se superpone ni entrecruza, con las descripciones extensionales. Es decir, dicha explicación no revela rasgos adicionales de los que puedan finalmente describirse extensionalmente. Su aporte consiste justamente en describir de modo diferente los mismos procesos cognitivos. En este sentido, todo el valor de las explicaciones intencionales es heurístico, lo cual no es poco en absoluto: el aporte heurístico del vocabulario mentalista es indispensable y no negociable para el estudio de las operaciones mentales.

De este modo, en lo que respecta a su relación con las disciplinas científicas, el nivel explicativo intencional es concebido estrictamente como un *modo de abordaje* de los problemas a los que aquellas deberán hacer frente: esto es, uno entre otros modos de abordaje, que es sin embargo ineludible dada la contribución heurística única y de otra manera inalcanzable que otorga. Ahora, sigue siendo cierto que las explicaciones y caracterizaciones intencionales no explican y caracterizan su objeto de forma directa. Aquellas son equiparadas por Dennett con “préstamos de inteligencia”<sup>181</sup>, préstamos de los que el teórico puede tomar provecho atribuyendo un contenido determinado a cualquier estado o proceso, o cualquier fragmento de ellos, que de este modo es puesto en relación con otros. Son préstamos en tanto que son usados y abandonados en cuanto hayan cumplido su rol. Pero, sin esta forma de abordar el problema de la descripción de las operaciones mentales, el teórico es dejado sin modo alguno de referirse a la función mental (o cerebral).

---

<sup>179</sup> Dennett (1969, p. 112).

<sup>180</sup> Quine (1960, p. 219).

<sup>181</sup> Dennett (1978a, p. 12). La traducción es mía.

El camino explicativo que vincula el contenido semántico de un estado intencional, por un lado, con sus mecanismos efectivos de implementación y la estructura en que se revelan, por otro, es entonces irreduciblemente indirecto. Para Dennett, éste es un hecho incontestable:

Lo que hace posible en última instancia que algo en este sentido represente algo dentro de un sistema es que tenga una función dentro de ese sistema [...]. Luego, el contenido es una función de la función, pero no cualquier estructura puede realizar cualquier función, puede confiadamente garantizar las relaciones normales requeridas. Entonces la función es una función de la estructura. Hay, luego, fuertes limitaciones estructurales indirectas sobre las cosas a las que podemos dotar de contenido<sup>182</sup>.

El principal error de Fodor es justamente suponer que pueda plantearse, de algún modo justificado, un vínculo directo entre contenido y estructura. Y, más aún, que el estudio científico de los procesos mentales, la psicología cognitiva propiamente dicha, pueda y deba modelarse sobre este supuesto.

En una primera aproximación, pareciera que esta perspectiva fodoriana es clara, simple y con repercusiones directas y evidentes para la investigación psicológica. Por otro lado, la relación entre los niveles (personal y sub-personal) de explicación es, sin lugar a duda, uno de los puntos más criticados de la teoría de Dennett. Puede, sin embargo, darse una explicación positiva de este contraste, una explicación que se refleja en los sentidos en que, al final del apartado 4.3, caracterizaba la teoría fodoriana de idealizada y anti-biológica frente a la teoría evolucionista dennettiana.

La medida en que la teoría representacional de la mente, de acuerdo a dicha caracterización, incorporaba una visión simplista de los problemas acerca de la intencionalidad y de su realidad psicológica, se transfiere a lo que idealmente sería el formato de las disciplinas que pretendan resolver aquellos problemas. Me refiero, fundamentalmente, a los aspectos de la teoría según los cuales el *sentido mismo* de las atribuciones intencionales del tipo de las que todos hacemos cotidianamente queda finalmente anclado en una realidad psíquica absolutamente correspondiente (lo que llamaba el “logicismo” fodoriano).

Una de las consecuencias de esta manera de ver la problemática concerniente la intencionalidad y los estados mentales es la concentración casi excluyente, por parte de Fodor, en el tipo de explicación propia de la psicología cognitiva (o lo que Fodor

---

<sup>182</sup> Dennett (1978a, p. 106). La traducción es mía.

interpreta como tal). En el fondo, la teoría representacional de la mente, con su apelación fuerte a representaciones mentales y operaciones computacionales, propone, entre otras cosas, *el* esquema básico para *la* solución de dicha problemática. Este esquema y esta solución refieren a una sola disciplina (y a un solo formato de ella), la psicología cognitiva, claramente definida en sus métodos y objetivos.

Otra consecuencia importante de mantener una relación directa entre estructura y contenido es que se debilita la distinción útil, casi unívocamente aceptada<sup>183</sup> (y que, por otra parte, para el caso de Dennett se refuerza), entre un nivel personal y un nivel sub-personal de explicación. Fodor, de hecho, llega a negar explícitamente que la distinción sea en absoluto siquiera relevante para la psicología cognitiva:

[C]ualquiera que sea la importancia que pueda tener para *ciertos* objetivos la distinción entre los estados del organismo y los estados de su sistema nervioso, no hay ninguna razón especial para suponer que es importante para los objetivos de la psicología cognitiva. [...] Si se tienen presentes [sus] objetivos, resulta [...] que la distinción ordinaria entre lo que hace, conoce, piensa y sueña el organismo y lo que ocurre a y en su sistema nervioso, no parece tener una importancia excesiva<sup>184</sup>.

La situación de la teoría de los sistemas intencionales es esencialmente inversa a la hasta aquí descrita. Aunque la teoría se apoya fuertemente en la distinción personal / sub-personal, no se *compromete* con ninguna disciplina en particular o alguna versión de ella: se limita a favorecer o sólo subrayar el rol especialmente relevante que algunos emprendimientos teóricos pueden tener, dados ciertos objetivos. La teoría tampoco identifica de modo claro un conjunto definitivo de objetivos a conseguir. En general, puede decirse que Dennett se compromete con menos, incluso sólo metodológicamente.

Creo que en general esta actitud es positiva y responde, entre otras cosas, a una comprensión de la complejidad del conjunto de fenómenos del que se trata. Esta comprensión está profundamente arraigada en el pensamiento dennettiano en tanto está, a la vez, íntimamente relacionada con el núcleo del enfoque evolucionista en el que su teoría se encuadra: una comprensión de la mente en tanto que inmersa en el mundo donde se despliegan sus operaciones; una comprensión por la cual las creencias y los deseos son entendidos, fundamentalmente, en términos de su contribución efectiva a la

---

<sup>183</sup> Aunque también hay quienes actualmente la critican. Ver, por ejemplo, Bermúdez (2000).

<sup>184</sup> Fodor (1975, p. 71).

regulación de la conducta de los organismos, y no, fundamentalmente, como aquellas cosas de las que puede decirse que son verdaderas o falsas, satisfechas o frustradas<sup>185</sup>.

Dos ejemplos particularmente interesantes de programas de investigación favorecidos por la teoría de los sistemas intencionales son el programa de la inteligencia artificial (o IA) y la etología cognitiva. Estas dos disciplinas son los puntos de referencia extra-filosófica más recurrentes en la obra de Dennett (ver, en especial, *Brainchildren*), y ambas revelan aspectos centrales netamente evolucionistas.

El interés en los resultados de la etología cognitiva refleja la importancia que tiene para Dennett el estudio de los sistemas intencionales no lingüísticos para poder abordar las cuestiones más intrincadas respecto de nuestro propio caso. Además de aclarar el campo de las complicaciones teóricas y metodológicas generadas por nuestra inmersión en el lenguaje, el estudio de organismos no lingüísticos, y sobre todo el estudio de los organismos evolutivamente más cercanos al ser humano, brinda una valiosísima perspectiva comparativa e incluso genética sobre nuestra conducta y los mecanismos que la controlan. Por otra parte, el estudio de sistemas intencionales simples proporciona un asiento empírico para dirimir cuestiones más puramente filosóficas acerca de temas generales como, por ejemplo, el de la inteligencia y el de la conciencia.

Dennett es un convencido partidario y defensor de la IA. Creo que esto debe destacarse en cuanto que el evolucionismo dennettiano encuentra respaldo en el trabajo de ingeniería inversa de las investigaciones en IA. Don Ross, comentando a Dennett, nota una afinidad de principio entre ambos enfoques teóricos:

Dado que la selección natural no puede volver atrás y rediseñar un diseño, una vez que las limitaciones ambientales y morfológicas la llevan a seguir una trayectoria determinada, el espacio disponible en las posibilidades de diseño se encoge de un modo lo suficientemente ajustado como para permitir que la ingeniería inversa genere hipótesis precisas y fundamentadas<sup>186</sup>.

En segundo lugar, este tipo de investigación se asienta en una descripción detallada de la *conducta particular* de un sistema intencional; el trabajo de ingeniería se emprende independientemente del estudio de las complejidades estructurales conocidas de algún sistema intencional determinado, atendiendo únicamente a la especificación de un desempeño conductual puntual. En tercer lugar, una ventaja de esta estrategia es que

---

<sup>185</sup> Pinker (2005, pp. 18-19) expone una crítica muy similar en espíritu a la que hago aquí.

<sup>186</sup> Ross, D., Brook, A., Thompson, D. (eds.), (2000, p. 4). La traducción es mía.

permite que se traigan a luz las características más “banales” (pero importantísimas) de la cognición, las que pasan desapercibidas por su obviedad e inespecificidad (y las que una perspectiva intelectualista típicamente suele pasar por alto)<sup>187</sup>.

Un aspecto fundamental del modo en que Dennett visualiza el desarrollo del estudio científico sobre la mente es lo que llama propiamente “la teoría del sistema intencional” (Dennett, 1987a, pp. 62-65). Esta teoría incluye un trabajo de, digamos, limpieza conceptual y de “legitimación” del vocabulario intencional, y consiste en la especificación detallada del perfil de un sistema cognitivo, para una tarea cognitiva particular, desde la actitud intencional; en otras palabras, se intenta especificar (a nivel personal) toda la información que el sistema debiera idealmente usar, recibir o transmitir para desarrollar alguna tarea determinada. Para esto, se hace uso de los términos mentalistas, pero se los intenta sistematizar y definir para que adquieran una aplicación técnica, relativamente formal, dentro de la teoría.

Dennett entiende esta tarea como parte integrante tanto de los estudios en etología cognitiva (Dennett, 1987a, p. 227) como de los programas de IA (Dennett, 1987a, p. 76). Las atribuciones intencionales que allí se especifican ponen en relación los agentes, interpretados como sistemas unitarios o totales, con el entorno en que interactúan. Ahora bien, queda claro que los modelos de IA, y especialmente los modelos conexionistas favorecidos por Dennett, no mantienen una relación semántica directa (de implementación o de algún tipo de traducción, por ejemplo) respecto de los acontecimientos del mundo externo: esto es, se preserva la distinción entre la explicación a nivel personal y sub-personal.

El nivel personal de explicación también se despliega en otro sentido, afín al de la teoría del sistema intencional, tal como la he caracterizado. Por ejemplo, en el caso de los modelos de IA:

Parte del atractivo de IA es que proporciona un modo de empezar por lo que son esencialmente categorías y diferencias fenomenológicas – características de los mundos nocionales– y retroceder hasta hipótesis sobre cómo implementar esas aptitudes<sup>188</sup>.

El atractivo al que el autor se refiere es propio de un emprendimiento teórico, coincidente o estrechamente relacionado con la teoría del sistema intencional, que

---

<sup>187</sup> El conocido “problema del marco” en IA es justamente el problema de dar cuenta de la enorme cantidad de información que un sistema cualquiera debe de algún modo manejar para desempeñar las tareas más simples.

<sup>188</sup> Dennett (1987a, p. 152).

Dennett ha denominado “heterofenomenología”<sup>189</sup>. Esta consiste en la caracterización del “mundo nocional” de un sistema intencional.

La idea de un mundo nocional es esencialmente la idea, conocida desde la fenomenología husserliana, del conjunto de objetos intencionales, ideas o representaciones (en sentido laxo) propios de la vida mental de un individuo. Sin embargo, el método para determinar este conjunto de objetos no es subjetivo o introspectivo (esto es, por *epoché* o reducción fenomenológica), sino que es puesto en práctica estrictamente desde una perspectiva de tercera persona, que sólo apela a hechos públicos, no privados.

Mi interés en traer a colación el método heterofenomenológico está en subrayar cómo lo que describe (al igual que en el caso de la teoría del sistema intencional) solamente se fija a partir de la observación, desde la perspectiva de la tercera persona, de hechos objetivos y públicamente accesibles respecto del desempeño efectivo de un agente: es decir, fundamentalmente, de las capacidades y disposiciones del sistema intencional, inserto en el medio en que se encuentra. En definitiva, lo que llamé la “primacía de la conducta” en la teoría dennettiana tiene en este sentido una repercusión meta-teórica importante<sup>190</sup>.

Volviendo rápidamente a Fodor, y descontando su ya advertida indiferencia hacia las investigaciones sobre el comportamiento animal en general, vale la pena notar que éste mantiene una postura crítica, negativa, respecto de la importancia del enfoque de la IA. Defendiendo la inviabilidad del proyecto de IA en sus primeras encarnaciones, Fodor sostiene que la razón de esto no es “porque tiene la imagen equivocada de la mente, sino porque tiene una mala metodología para transformar esa imagen en ciencia. No creo que uno hace ciencia sobre fenómenos complejos tratando de construir modelos de alguna variación [*variance*] gruesa observable”<sup>191</sup>. Una vez más, creo que esta oposición hacia los métodos de la IA puede sencillamente explicarse teniendo en cuenta su estrecha relación con el comportamiento efectivo y, por otro lado, el tajante distanciamiento respecto de la conducta de los organismos que, en el previo apartado, atribuimos a la teoría representacional de la mente y, especialmente, a la metodología que propone para la psicología.

---

<sup>189</sup> Dennett introduce este enfoque (aunque sin bautizarlo) en “Two Approaches to Mental Images” (en Dennett, 1978a).

<sup>190</sup> También es relevante mencionar cómo el cometido, filosófico y meta-teórico, de la configuración de diferentes “tipos de mentes” (destacado en p. 59 a raíz del par creencia / opinión) puede beneficiarse tanto de la teoría del sistema intencional como de la heterofenomenología.

<sup>191</sup> En Loewer y Rey (eds.), (1991, p. 279).

Una última observación epistemológica. Hay una diferencia crucial respecto de la idea que Dennett y Fodor tienen de la psicología en tanto que ciencia. En la medida que la psicología sea, como ambos filósofos aceptan, mentalista en cuanto al vocabulario en el que se desarrolla, esto la lleva, en opinión de Dennett, al nivel de una ciencia como la biología (interpretada en su versión adaptacionista). El adaptacionismo en biología se corresponde al mentalismo en psicología en tanto supone la incorporación de un nivel descriptivo y explicativo que hace sistemáticamente imposible la obtención de “leyes de la naturaleza”, en el sentido en que lo haría la física, por ejemplo. La neutralidad respecto de la implementación de los estados caracterizados intencionalmente responde al hecho aceptado de que, en general, los problemas acerca de la intencionalidad no pueden ser resueltos desde una perspectiva microrreductiva, sino apelando a una perspectiva que va desde los fenómenos “gruesos” hasta gradualmente llegar a los procesos específicos detrás de ellos<sup>192</sup>. Esta perspectiva es un modelo entre otros de investigación científica. Por otra parte, a los ojos de Fodor la búsqueda de “leyes” psicológicas directamente compromete el estatus de la psicología en tanto que ciencia (“...si no hay leyes intencionales, entonces no se puede hacer ciencia a partir de explicaciones intencionales”<sup>193</sup>).

---

---

<sup>192</sup> Un aspecto filosóficamente interesante de esto es que, en cuanto mantengamos que contenido y estructura no pueden ser relacionados de modo directo, el llamado “problema mente / cuerpo” es, en el fondo, un problema mal planteado, por lo que justamente supone esta relación directa.

<sup>193</sup> Fodor (1998, p. 24).

## Conclusión

Quisiera, antes de hacer algunas reflexiones finales, mencionar algunos temas muy relevantes en el pensamiento dennettiano, que me hubiera gustado desarrollar, pero que, por una razón u otra, no tuvieron espacio en este trabajo. Como sostuve en el capítulo 3, el concepto de “opinión” es central para dar cuenta del modo cómo el lenguaje tiene un efecto retrospectivo que altera profundamente nuestra particular vida mental; en esta dirección, otro concepto central, tomado de las contribuciones de Richard Dawkins (1976), es el de “meme” y, con él, lo que Dennett llama “memética”, una propuesta para un estudio posible de los modos de “transmisión” cultural. Hay, por otra parte, una importante distinción dennettiana (1991a, pp. 316 y siguientes) entre “reportar”, a través de contenidos lingüísticos, y “expresar”, a través del comportamiento, los propios estados intencionales, que apunta en el sentido de determinar la medida en que los criterios de atribución intencional no dependen del repertorio conceptual del agente interpretado. Otra línea de investigación que quisiera destacar, especialmente por su relevancia para la cuestión del evolucionismo, es lo que se conoce como “cognición corporizada” (*embodied cognition*), que intenta estudiar los sentidos en que mantener una separación tajante entre mente, por un lado, y cuerpo, acción y mundo, por otro, es altamente contraproducente.

Recapitulando. En este trabajo he tratado de mostrar que hay un desencuentro profundo entre las teorías que Daniel Dennett y Jerry Fodor proponen para dar cuenta, en general, de la intencionalidad de los estados mentales, un desencuentro que abarca tanto el nivel de las discusiones y los argumentos particulares como el de las estrategias de abordaje y las maneras de enfocar los problemas de una teoría del contenido mental: en definitiva, un desencuentro profundo entre dos maneras de configurar las cuestiones y los problemas de los que se ocupa (o mejor, de los que se debería ocupar) una teoría filosófica de la mente y que, en este sentido, puede concebirse finalmente como una incompatibilidad de raíz en la articulación del concepto general de “mente” en el que las teorías dennettiana y fodoriana, de algún modo, se asientan.

He proporcionado argumentos para justificar que esta incompatibilidad puede explicarse apelando al modelo teórico y metodológico proporcionado por la teoría

darwiniana de la evolución por selección natural: en la medida en que las teorías de la intencionalidad estudiadas incorporen o no dicho modelo, puede darse cuenta plenamente de la distancia entre un planteo gradualista y contextualista como el dennettiano y un planteo discontinuista e intelectualista como el fodoriano. La resultante diferencia de base para una concepción general de la mente y el mismo concepto asumido de “mente”, es la que hay entre interpretarla aisladamente, como el ámbito operacional articulado que de manera unidireccional regula la conducta inteligente, e interpretarla no aisladamente, como sumida en el contexto real donde se inserta, contexto que incluye el organismo “total”, su conducta y el mundo que manipula (una diferencia reflejada en parte por el contraste, propuesto en el apartado 4.3, entre la primacía de la conducta y la primacía de la conciencia)<sup>194</sup>.

He argumentado, en particular, que el alcance del enfoque evolucionista acerca del contenido mental puede mostrarse apelando a la fundamental distinción dennettiana entre creencias y opiniones, distinción que principalmente pone en relieve la importancia de mantener separados el ámbito de la intencionalidad del de la competencia lingüística. Esta justificación se despliega sobre los diferentes aspectos centrales, teóricos y extra-teóricos, de las posiciones filosóficas estudiadas. Estos comprendieron: el asiento teórico o punto de partida de estas posiciones, esto es, según sostuve, su interpretación de la psicología de sentido común; su metodología, cuyo estudio incorporó de modo central una crítica radical a la noción de “actitud proposicional”; sus tesis sustantivas, especialmente las referidas a la configuración de los estados intencionales y la determinación de su contenido; y, finalmente, su aspecto meta-teórico general. (Estos puntos se corresponden con los distintos apartados del capítulo 4).

Teniendo en consideración la distinción en la que se basó mi argumentación, puede afirmarse que el enfoque anti-evolucionista acerca del contenido mental, con toda la profundidad de sus consecuencias, depende fuertemente de una visión restrictiva de la atribución mental que surge de una analogía implícita con el modelo de nuestro lenguaje natural. A su vez, el enfoque evolucionista acerca del contenido mental se asienta en general en la constatación de un hecho crucial para una teoría de la intencionalidad: el hecho de que el lenguaje y la competencia lingüística deben ubicarse en el contexto más amplio de la atribución intencional. El lenguaje humano es, sin lugar a dudas, el medio

---

<sup>194</sup> Me parece interesante notar el mayor y menor grado en que, respectivamente, las visiones fodoriana y dennettiana permiten dar una “definición” de la mente.

por excelencia de atribución de contenido intencional. Aún así, una visión evolucionista acerca de la intencionalidad permite identificarlo plenamente como medio, esto es, como una herramienta particular con características particulares que configuran el modo en que nosotros abordamos el mundo.

La importancia del evolucionismo en el ámbito estudiado no significa que toda vez que queramos hablar sobre la mente tengamos que hablar sobre Darwin. Más bien, el impacto filosófico del evolucionismo implica que toda vez que queramos hablar sobre la mente tengamos que hablar sobre el modo en que hablamos sobre la mente o, menos circularmente, acerca del modo cómo conocemos la mente. Esta es, en el fondo, la importancia de mantener una *perspectiva* evolucionista, como diferente de una *postura* evolucionista (que, en algunos aspectos, es también importante y, a veces, imprescindible), para todos los aspectos del estudio de la mente.

El carácter fuertemente general, propiamente “filosófico”, que he mantenido en este trabajo responde a la que me parecía una necesidad de poner en claro algunos desencuentros de base en una serie de disputas improductivas entre (sin lugar a dudas) dos de los filósofos contemporáneos de la mente más influyentes. Puntualmente, dicho carácter me pareció justificado en la idea de que una gran cantidad de discusiones mal planteadas, vacías, podrían evitarse, si se aclarara el marco teórico y metodológico en que la investigación filosófica y no filosófica se embarca (de aquí, también, el carácter predominantemente crítico de mis aportes). Si existe un nivel de contribución (de nuevo) más propiamente *filosófico*, me da la impresión de que, en un espíritu afín (pero no equivalente) al del viejo modo wittgensteineano, es éste.

Esto no significa que no haya otros rumbos posibles para el aporte filosófico, que puedan llegar a ser importantes dentro y fuera de la disciplina: otros rumbos que, en particular, se aproximen más a las investigaciones empíricas o, quizás sería mejor decir, aplicadas, y que de este modo puedan considerarse más o menos naturalizados, quineanamente hablando. Pero ya notamos las diferencias de *tono* entre nuestros autores. Específicamente, ya notamos el sentido en que, según lo defendido, un aporte como el de Dennett es, por lo menos, más “confiable” que el modo en que Fodor contribuye, o pretende contribuir, de forma directa a la psicología (un estilo que él mismo denominó “psicología especulativa”<sup>195</sup>). Mientras éste trata de elaborar un modelo, aquel trata de elaborar una perspectiva, en especial, intentando proporcionar

---

<sup>195</sup> Fodor (1975, p. 16).

herramientas (metáforas, comparaciones, ejemplos, experimentos mentales, etcétera) para pensar mejor y, sobre todo, para encarar mejor la investigación, del tipo que sea.

En este sentido, puede vislumbrarse en el estilo dennettiano un camino entre Wittgenstein y Quine, que se refleja en la medida (importante, en mi opinión) en que el estudio aplicado (y no aplicado), empírico (y no empírico) de la mente puede beneficiarse del esclarecimiento conceptual, de la problematización, de la formulación de buenos cuestionamientos y, por qué negarlo, de la mera fantasía. A fin de cuentas, lo que es más importante, y lo que el evolucionismo *como metáfora* debe sugerirnos a los filósofos, es que abrir caminos, mantener una perspectiva abierta y flexible sobre los problemas, es siempre preferible a cerrarlos.

---

## Bibliografía

- Allen, C. (1992), "Mental Content", *British Journal for the Philosophy of Science*, 43, pp. 537-553.
- Ben-Yami, H. (1997), "Against Characterizing Mental States as Propositional Attitudes", *Philosophical Quarterly*, 186, pp. 84-89.
- Bennett, J., (1993), "Comments on Dennett form a Cautious Ally", *Behavioral and Brains Sciences*, 16, pp. 381-416.
- Bermúdez, J. L. (2000), "Personal and Sub-Personal: A Difference without a Distincion", *Philosophical Explorations*, 3 (1), pp. 63-82.
- (2003), *Thinking without Words*, Oxford University Press, Oxford.
- Bogdan, R. (1997), *Interpreting Minds*, MIT Press, Cambridge, MA.
- Brook, A., Ross, D. (eds.), (2002), *Daniel Dennett (Contemporary Philosophy in Focus)*, Cambridge University Press, Nueva York.
- Burge, T. (1977), "Belief *de re*", *Journal of Philosophy*, 76 (6), pp. 338-362.
- (1986), "Individualism and Psychology", *Philosophical Review*, 95 (1), pp. 3-45.
- Carruthers, P., Boucher, J. (eds.), (1998), *Language and Thought*, Cambridge University Press, Cambridge.
- Churchland, P. M. (1981), "Eliminative Materialism and the Propositional Attitudes", *Journal of Philosophy*, 78 (2), pp. 67-90.
- Clark, A. (1988), "Thoughts, Sentences and Cognitive Science", *Philosophical Psychology*, 1, pp. 263-278.
- (1990), "Belief, Opinion and Consciousness", *Philosophical Psychology*, 3 (1), pp. 139-154.

- (1996), *Being There*, MIT Press, Cambridge, MA.
- (2002), “Minds, Brains and Tools”, en Clapin, H. (ed.), *Philosophy of Mental Representation*, Clarendon Press, Oxford, pp. 66-90.
- Clark, A., Chalmers, D. (1998), “The Extended Mind”, *Analysis*, 58, pp. 10-23.
- Clark, A., Karmiloff-Smith, A. (1993a), “The Cognizer’s Innards”, *Mind and Language*, 8 (4), pp. 487-519.
- (1993b), “What’s Special about the Development of the Human Mind / Brain?”, *Mind and Language*, 8 (4), pp. 569-581.
- Dahlbom, B. (ed.), (1993), *Dennett and His Critics*, Blackwell, Oxford.
- Davidson, D. (1975), “Pensamiento y Habla”, en Guttenplan, S. (ed.), *Mind and Language*, Clarendon Press, Oxford, pp. 7-23. Reimpreso en *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford, 1984 [*De la Verdad y de la Interpretación*, Gedisa, Barcelona, 1990].
- (1982), “Rational Animals”, *Dialectica*, 36 (4), pp. 318-327.
- De Waal (1996), *Good Natured*, Harvard University Press, Cambridge [*Bien Natural*, Herder, Barcelona, 1997].
- Dennett, D. (1969), *Content and Consciousness*, Routledge and Kegan Paul, Londres [*Contenido y Conciencia*, Gedisa, Barcelona, 1996].
- (1971), “Intentional Systems”, *Journal of Philosophy*, 68, pp. 87-106. Reimpreso en Dennett (1978a).
- (1975), “Why the Law of Effect Will Not Go Away”, *Journal for the Theory of Social Behaviour*, 5, pp. 169-87. Reimpreso en Dennett (1978a).
- (1977), “A Cure for the Common Code?”, *Mind*, 86, pp. 265-80. Reimpreso en Dennett (1978a).
- (1978a), *Brainstorms*, MIT Press, Cambridge, MA.

- (1978b), “Beliefs about Beliefs”, *Behavioral and Brain Sciences*, 1 (4), pp. 568-570.
- (1981a), “True Believers”, en Heath, A. (ed.), *Scientific Explanations*, Oxford University Press, Oxford, pp. 53-75. Reimpreso en Dennett (1987).
- (1981b), “Three Kinds of Intentional Psychology”, en Healey, R. (ed.), *Reduction, Time and Reality*, Cambridge University Press, Cambridge, pp. 37-61. Reimpreso en Dennett (1987).
- (1982a), “How to Study Human Consciousness Empirically”, *Synthese*, 53 (2), pp. 159-180.
- (1982b), “Beyond Belief”, en Woodfield, A. (ed.), *Thought and Object*, Clarendon Press, Oxford, pp. 1-96. Reimpreso en Dennett (1987).
- (1987a), *The Intentional Stance*, MIT Press, Cambridge, MA [*La Actitud Intencional*, Gedisa, Barcelona, 1998].
- (1987b), “Evolution, Error and Intentionality”, en Wilks, Y., Partridge, D. (eds.), *Sourcebook on the Foundations of Artificial Intelligence*, New Mexico University Press, pp. 190-211. Reimpreso en Dennett (1987).
- (1988a), “Précis of *The Intentional Stance*”, *Behavioral and Brain Sciences*, 11 (3), pp. 495-546.
- (1988b), “Review of J. Fodor, *Psychosemantics*”, *Journal of Philosophy*, 85, pp. 384-389.
- (1990), “The Interpretation of Texts, People and Other Artifacts”, *Philosophy and Phenomenological Research*, 50, pp. 177-194.
- (1991a), *Consciousness Explained*, Little, Brown and Company, Nueva York [*La Conciencia Explicada*, Paidós, Buenos Aires, 1995].

- (1991b), “Mother Nature versus the Walking Encyclopedia”, en Ramsey, W., Stich, S., Rumelhart, D. (eds.), *Philosophy and Connectionist Theory*, Erlbaum, Hillsdale, pp. 21-30.
- (1991c), “Real Patterns”, *Journal of Philosophy*, 88 (1), pp. 27-51. Reimpreso en Dennett (1998a).
- (1991d), “Two Contrasts: Folk Craft versus Folk Science, and Belief versus Opinion”, en Greenwood, J. (ed.), *The Future of Folk Psychology*, Cambridge University Press, Cambridge, pp. 135-148. Reimpreso en Dennett (1998a).
- (1993), “Learning and Labeling”, *Mind and Language*, 8 (4), pp. 540-548.
- (1994), “The Role of Language in Intelligence”, en Khalifa, J. (ed.), *What is Intelligence?*, Cambridge University Press, Cambridge, pp. 161-178.
- (1995), *Darwin’s Dangerous Idea*, Simon and Schuster, Nueva York.
- (1996a), *Kinds of Minds*, Weidenfeld & Nicolson, Londres [*Tipos de Mentes*, Debate, Madrid, 2000].
- (1996b), “Response to Fodor on *Darwin’s Dangerous Idea*”, *Mind and Language*, 113, pp. 263-269.
- (1998a), *Brainchildren*, MIT Press, Cambridge, MA.
- (1998b), “The Myth of Double Transduction”, en Hameroff, S. (ed.), *Toward a Science of Consciousness II*, MIT Press, Cambridge, MA, pp. 97-107.
- (1998c), “An Overview of My Work in Philosophy”, *Contemporary British and American Philosophy and Philosophers*, Kang, O., Fuller, S. (eds.), People's Press, Beijing.
- (2000), “Making Tools for Thinking”, en Sperber, D. (ed.), *Metarepresentations*, Oxford University Press, Nueva York, pp. 17-29.

- (2001), “Things about Things”, en Branquinho, J. (ed.), *The Foundations of Cognitive Science*, Clarendon Press, Oxford, pp. 133-143.
- (2002), “Reply to Clark”, en Clapin, H. (ed.), *Philosophy of Mental Representation*, Clarendon Press, Oxford, pp. 91-93.
- Dennett, D., Haugeland, J. (1987), “Intentionality”, en Gregory, R. (ed.), *The Oxford Companion to the Mind*, Oxford University Press, Oxford, pp. 383-386.
- Dewey, J. (1910), “The Influence of Darwinism on Philosophy”, en Dewey, J., *The Influence of Darwinism on Philosophy and Other Essays in Contemporary Thought*, Henry Holt and Co., Nueva York, pp. 1-19 [“La Influencia del Darwinismo en la Filosofía”, en Faerna, A. (ed.), *La Miseria de la Epistemología*, Biblioteca Nueva, Madrid, 2000, pp. 49-60].
- Dretske, F. (1981), *Knowledge and the Flow of Information*, MIT Press, Cambridge, MA [Conocimiento e Información, Salvat, Barcelona, 1987].
- Fodor, J. (1975), *The Language of Thought*, Harper and Row, Nueva York [El Lenguaje del Pensamiento, Alianza, Madrid, 1984].
- (1978), “Propositional Attitudes”, *The Monist*, 61 (4), pp. 501-523. Reimpreso en Fodor (1981) [“Las Actitudes Proposicionales”, en Rabossi, E. (ed.), *Filosofía de la Mente y Ciencia Cognitiva*, Paidós, Barcelona, 1995, pp. 173-203].
- (1980), “Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology”, *Behavioral and Brain Sciences*, 3, pp. 63-109. Reimpreso en Fodor (1981).
- (1981), *Representations*, MIT Press, Cambridge, MA.
- (1983), *The Modularity of Mind*, MIT Press, Cambridge, MA.
- (1985), “Fodor’s Guide to Mental Representation”, *Mind*, 94, pp. 76-100.

- (1986), “Why Paramecia Don’t Have Mental Representations”, *Midwest Studies in Philosophy*, 10, pp. 3-23.
- (1987), *Psychosemantics*, MIT Press, Cambridge, MA [*Psicosemántica*, Tecnos, Madrid, 1994].
- (1990), “A Theory of Content II: The Theory”, en Fodor, J., *A Theory of Content and Other Essays*, MIT Press, Cambridge, MA, pp. 89-136.
- (1994), *The Elm and the Expert*, MIT Press, Cambridge, MA.
- (1998), *Concepts*, Clarendon Press, Oxford [*Conceptos*, Gedisa, Barcelona, 1999].
- (2004), “Having Concepts”, *Mind and Language*, 19 (1), pp. 29-47.
- Fodor, J., McLaughlin, B. (1998), “Connectionism and the Problem of Systematicity (Continued)”, *Cognition*, 62, pp. 109-119.
- Föllesdal, D. (1990), “Indeterminacy and Mental States”, en Barrett, R., Gibson, R. (eds.), *Perspectives on Quine*, Blackwell, Oxford, pp. 98-109.
- Foss, J. (1994), “On the Evolution of Intentionality as Seen from the Intentional Stance”, *Inquiry*, 37 (3), pp. 287-310.
- Frankish, K. (1998), “A Matter of Opinion”, *Philosophical Psychology*, 11, pp. 423-442.
- Jackendoff, R. (1992), *Languages of the Mind*, MIT Press, Cambridge, MA.
- Johansson, S. (2001), *Animal Communication, Animal Minds, and Animal Language*, Bachelor’s Thesis, University of Lund.
- Loar, B. (1983), “Must Beliefs Be Sentences?”, en Asquith, P., Nickles, T. (eds.), *Proceedings of the Philosophy of Science Association for 1982*, East Lansing, Michigan.
- Loewer, B., Rey, G. (eds.), (1991), *Meaning in Mind*, Blackwell, Oxford.

- Lycan, W. (1990), "Mental Content in Linguistic Form", *Philosophical Studies*, 58, pp. 147-154.
- Papafragou, A. (2002), "Mindreading and Verbal Communication", *Mind and Language*, 17 (1, 2), pp. 55-67.
- Pinker, S. (1994), *The Language Instinct*, Penguin, Harmondsworth [*El Instinto del Lenguaje*, Alianza, Madrid, 2001].
- (2005), "So How *Does* the Mind Work?", *Mind and Language*, 20 (1), pp. 1-24.
- Pinker, S., Bloom, P. (1990), "Natural Language and Natural Selection", *Behavioral and Brain Sciences*, 13, pp. 707-784.
- Premack, D., Woodruff, G. (1978), "Does the Chimpanzee Have a Theory of Mind?", *Behavioral and Brain Sciences*, 1, pp. 515-526.
- Preston, J. (ed.), (1997), *Thought and Language*, Cambridge University Press, Nueva York.
- Preston, S., de Waal, F. (2002), "Empathy: Its Ultimate and Proximate Bases", *Behavioral and Brain Sciences*, 25, pp. 1-20.
- Prinz, J., Clark, A. (2004), "Putting Concepts to Work", *Mind and Language*, 19 (1), pp. 56-69.
- Putnam, H. (1975), "The Meaning of 'Meaning'", en Gunderson, K. (ed.), *Language, Mind and Knowledge*, University of Minnesota Press, Minneapolis, pp. 131-193 ["El Significado de 'Significado'", *Crítica*, Ciudad de México, 1984].
- Quine, W. V. O. (1960), *Word and Object*, MIT Press, Cambridge, MA [*Palabra y Objeto*, Labor, Barcelona, 1968].
- (1969), "Propositional Objects", en Quine, W. V. O., *Ontological Relativity and Other Essays*, Columbia University Press, Nueva York, pp. 139-160.

- Ross, D., Brook, A., Thompson, D. (eds.), (2000), *Dennett's Philosophy*, MIT Press, Cambridge, MA.
- Searle, J. (1980), "Minds, Brains and Programs", *Behavioral and Brain Sciences*, 3, pp. 417-424.
- (1983), *Intentionality*, Cambridge University Press, Nueva York.
- (2002), "Animal Minds", en Searle, J., *Consciousness and Language*, Cambridge University Press, Cambridge, pp. 61-76.
- Sedivy, S. (2004), "Minds: Contents without Vehicles", *Philosophical Psychology*, 17 (2), pp. 149-179.
- Skidelsky, L. (2003), *Representaciones Mentales*, manuscrito de tesis de doctorado.
- Sober, E. (1981), "The Evolution of Rationality", *Synthese*, 46, pp. 95-120.
- Stalnaker, R. (1990), "Mental Content and Linguistic Form", *Philosophical Studies*, 58, pp. 129-146.
- Stich, S. (1983), *From Folk Psychology to Cognitive Science*, MIT Press, Cambridge, MA.
- (1996), *Deconstructing the Mind*, Oxford University Press, Oxford.
- Tye, M. (1992), "Naturalism and the Mental", *Mind*, 101 (403), pp. 421-441.
- Varela, F., Thompson, E., Rosch, E. (1991), *The Embodied Mind*, MIT Press, Cambridge, MA.
- Westbury, C., Dennett, D., (2000), "Mining the Past to Construct the Future", en Schacter, D., Scarry, E. (eds.), *Memory, Brain, and Belief*, Harvard University Press, Cambridge, MA.
- Whiten, A. (1996), "Theory of Mind in Non-Verbal Apes", en Walsh, D. (ed.), *Naturalism, Evolution and Mind*, Cambridge University Press, Cambridge.

Wilkerson, W. (1999), "From Bodily Motions to Bodily Intentions", *Philosophical Psychology*, 12 (1), pp. 61-77.

Wittgenstein, L. (1953), *Philosophische Untersuchungen / Philosophical Investigations*, Blackwell, Oxford [*Investigaciones Filosóficas*, Crítica, Barcelona, 1988].

——— (1968), *The Blue and Brown Books*, Harper & Row, Nueva York [*Cuadernos Azul y Marrón*, Tecnos, Madrid, 1993].

——— (1969), *Über Gewissheit / On Certainty*, Blackwell, Oxford [*Sobre la Certeza*, Gedisa, Barcelona, 1988].

## Índice

|  |            |
|--|------------|
| <b>Introducción</b>  | <b>1</b>   |
| <br>   |            |
| <b>Capítulo 1: La Teoría de los Sistemas Intencionales</b>                       | <b>7</b>   |
| <i>1.1: Presentación General de la Teoría</i>                                    | 7          |
| <i>1.2: Carácter Gradualista de la Teoría</i>                                    | 13         |
| <i>1.3: Enfoque Evolucionista acerca del Contenido Mental</i>                    | 18         |
| <br>   |            |
| <b>Capítulo 2: La Teoría Representacional de la Mente</b>                        | <b>27</b>  |
| <i>2.1: Presentación General de la Teoría</i>                                    | 27         |
| <i>2.2: Enfoque Anti-Evolucionista acerca del Contenido Mental</i>               | 40         |
| <br>   |            |
| <b>Capítulo 3: La Distinción entre Creencia y Opinión</b>                        | <b>45</b>  |
| <i>3.1: Presentación y Análisis Preliminar de la Distinción</i>                  | 45         |
| <i>3.2: Carácter Gradualista de la Distinción</i>                                | 52         |
| <i>3.3: Relevancia e Implicaciones de la Distinción</i>                          | 60         |
| <br>   |            |
| <b>Capítulo 4: El Alcance del Evolucionismo</b>                                  | <b>69</b>  |
| <i>4.1: Concepciones Pragmatista e Intelectualista de la Psicología Popular</i>  | 69         |
| <i>4.2: Las Actitudes Proposicionales como Modelo de Análisis Filosófico</i>     | 82         |
| <i>4.3: La Primacía de la Conducta y el Realismo Intencional</i>                 | 96         |
| <i>4.4: Implicaciones Meta-Teóricas y la Relación entre Niveles Explicativos</i> | 115        |
| <br>   |            |
| <b>Conclusión</b>  | <b>126</b> |
| <br>   |            |
| <b>Bibliografía</b>  | <b>130</b> |