

# Representaciones mentales

## ¿Eslabón entre el individuo y el mundo?

Autor:

Skidelsky, Liza

Tutor:

Rabossi, Eduardo

2003

Tesis presentada con el fin de cumplimentar con los requisitos finales para la obtención del título en Doctor de la Universidad de Buenos Aires en Filosofía

Posgrado

TESIS 10-3-7

Universidad de Buenos Aires  
Facultad de Filosofía y Letras

FACULTAD de FILOSOFIA y LETRAS	
Nº 48740	MESA
29 AGO 2003	
Agr.	ENTRADA

# Representaciones Mentales: ¿Eslabón entre el individuo y el mundo?

## Tesis de Doctorado

UNIVERSIDAD DE BUENOS AIRES  
FACULTAD DE FILOSOFIA Y LETRAS  
Dirección de Bibliotecas

Doctoranda: Liza Skidelsky

Director: Eduardo Rabossi

Lugar: Buenos Aires

Fecha: Agosto 2003

# Índice

Prefacio .....	1
Abreviaturas y referencias .....	4
<i>Introducción</i>	
<i>Representaciones e intencionalidad</i> .....	5
<i>Capítulo 1</i>	
<i>La revolución cognitiva</i> .....	23
1. Ciencia cognitiva .....	24
1.1. El conductismo científico .....	25
1.2. Primera fuente: la reacción al conductismo .....	29
1.3. Segunda fuente: la inteligencia artificial .....	32
1.4. Ciencia cognitiva .....	33
1.5. Psicología cognitiva .....	40
1.5.1. Dos teorías de la psicología cognitiva .....	43
2. Filosofía de la psicología cognitiva .....	46
2.1. Intencionalidad .....	47
2.2. Funcionalismo filosófico .....	52
2.3. Funcionalismo computacional y psicología cognitiva .....	59
<i>Capítulo 2</i>	
<i>El problema de las representaciones y la teoría representacional de la mente</i> .....	63
1. El problema de las representaciones .....	63
1.1. Los criterios .....	65
1.2. Los requisitos .....	73
1.3. El formato .....	77
1.4. La implementación .....	78
1.5. La definición .....	80
2. La teoría representacional de la mente .....	82
2.1. La teoría .....	83
2.2. La teoría y el problema de las representaciones .....	92

*Capítulo 3*

<i>La semántica naturalista científicista</i> .....	98
1. La semántica informacional .....	101
2. La teleosemántica .....	109
3. El naturalismo científicista .....	121

*Capítulo 4*

<i>Naturalismo sin semántica intencional</i> .....	139
1. Críticas al naturalismo científicista .....	140
1.1. Primera estrategia: argumentos <i>a priori</i> .....	142
1.2. Segunda estrategia: críticas internas .....	144
1.3. Tercera y cuarta estrategia: hacia un naturalismo amplio .....	149
2. El naturalismo amplio .....	155

*Capítulo 5*

<i>Representaciones sin semántica intencional</i> .....	184
1. Estados subdoxásticos y estados doxásticos .....	186
1.1. Accesibilidad a la conciencia .....	188
1.2. Integración inferencial .....	193
1.3. Conceptualización .....	197
2. Contenido interno y contenido intencional .....	203
3. Contenido interno y contenido estrecho .....	216
4. Estados subpersonales y estados personales .....	223

*Conclusión*

<i>Representaciones e intencionalidad: computación y comunicación</i> .....	232
1. Representación .....	232
1.1. Cuestiones psicológicas y filosóficas .....	233
1.2. Representación y computación .....	237
1.3. Representación y comunicación .....	238
2. La distinción personal-subpersonal .....	241
3. Computación y comunicación .....	246
4. Representaciones mentales: ¿eslabón entre el individuo y el mundo? .....	248
Bibliografía .....	251

## Prefacio

Esta tesis de doctorado es el fruto de varios años de investigación y docencia en las áreas de metafísica, y filosofía de la mente y de la psicología y, de otros tantos años de investigación, docencia y trabajo clínico en el área de neuropsicología. Si bien la tesis se inscribe completamente en el ámbito de la filosofía de la psicología, mi doble formación en temas filosóficos y de la ciencia cognitiva me ha permitido enriquecer mis investigaciones aportándome una visión *in situ* de los problemas conceptuales y empíricos de la psicología cognitiva y de los aportes que la filosofía de la psicología puede hacer en relación con la clarificación de estos problemas.

Generalmente se piensa que aquellos filósofos que se ocupan de una disciplina científica deberían tener ciertos conocimientos teóricos y prácticos acerca de esa disciplina. Creo que esto es imprescindible para los que creen, como yo, en una filosofía naturalizada, esto es, una filosofía que sea continua con la ciencia de modo de desarrollarse en conformidad con, y hacer uso de, la ciencia corriente para contribuir a la explicación de los diferentes aspectos del ámbito de lo humano. Espero que esto no sea sólo un *desideratum* y pueda verse plasmado en esta tesis.

En este sentido, agradezco a aquellos que han contribuido a mi formación filosófica y científica. En primer lugar, a Eduardo Rabossi quien fue mi director de tesis y de las becas de investigación de iniciación y perfeccionamiento que he tenido en la Universidad de Buenos Aires durante el período 1995-2000, como así también director del grupo de investigación UBACyT de los martes, del que participo desde hace ya más de 10 años. Sin su guía filosófica no hubiera accedido a la comprensión de las discusiones en torno a temas de metafísica y filosofía de la mente. Su claridad conceptual y su constante estímulo han sido un motor invaluable de mi actividad filosófica. También quisiera agradecer a Antonio

Castorina, porque fue gracias a sus estimulantes clases sobre Chomsky y Piaget en un seminario de grado en la facultad de Filosofía y Letras que surgió en mí la pasión de la investigación que se cristalizó por primera vez en mi beca de estudiante (de la Universidad de Buenos Aires durante el período 1991-1993). Asimismo estoy en deuda con Diana Pérez con quien he discutido a lo largo de muchos años todos los temas que trato en mi tesis en innumerables charlas de café, almuerzos y en los grupos de investigación que dirige. Es un verdadero placer tener una colega amiga de la inteligencia, la paciencia y la amplia comprensión filosófica de Diana.

Quisiera agradecer también a la universidad de Buenos Aires por las becas que he tenido durante el período 1991-2000. Gracias a estas investigaciones he podido abordar un amplio espectro de temas, tales como el innatismo, la modularidad de la mente, las arquitecturas cognitivas en inteligencia artificial, y la intencionalidad, que enriquecieron mi comprensión de la psicología cognitiva y la filosofía de la psicología. También agradezco a Ernie Lepore por haberme permitido disfrutar del clima intelectual de la cuna de la filosofía de la psicología actual en el *Center for Cognitive Science* de *Rutgers University* durante el semestre de primavera de 2002. Haber asistido a los cursos de doctorado de J. Fodor, A. Goldman, E. Sosa, B. McLaughlin, y P. Pietroski sumado a los seminarios en los que participaron filósofos como S. Stich, B. Loar, C. y G. MacDonald y otros me ha permitido enriquecer mi reflexión filosófica de una manera sumamente fructífera. Asimismo agradezco a los miembros de los numerosos subsidios de investigación UBACyT en los que participo y he participado desde 1989, a los miembros de los grupos de investigación en los que he participado como invitada y a los alumnos, en especial de la materia Problemas Especiales de Metafísica donde siempre vemos temas de filosofía de la mente, quienes contribuyeron a la clarificación de muchas de las ideas que aparecen aquí expuestas.

Por último quisiera agradecer a mis colegas, incluso a algunos de ellos que no acuerdan con mucho de lo que sostengo, por discusiones filosóficas, en distintas circunstancias a lo largo de estos años, sin las cuales esta tesis se hubiera visto empobrecida: Julieta Arosteguy, Cristina González, Ana Hulton, Juan Rodríguez

Larreta, Sandra Lazzer, Hernán Miguel, Karina Pedace, Víctor Rodríguez, Pablo Rychter, Carolina Scotto, Rob Stainton, y Julia Vergara. Un agradecimiento muy especial al *lunch team* conformado por Sandra Lazzer, Diana Pérez y Julia Vergara con quienes he compartido discusiones encarnizadas, filosóficas y de las otras.

## Abreviaturas y referencias

El año de las citas y las referencias de los textos es el del original. Cuando en la bibliografía figura la reimpresión, la paginación se refiere a la misma. En el caso en que figure la traducción al español, la paginación es de la traducción.

Las siguientes abreviaturas escritas por orden alfabético aparecen en la tesis, entre paréntesis figura el capítulo en el que se introducen:

AP: actitud proposicional (cap. 1)

ES: estado subdoxástico (cap.5)

FF: forma fonética (cap. 1)

FL: forma lógica (cap. 1)

Lenguaje-I: lenguaje internalizado (cap. 1)

PNC: proyecto naturalista científicista (cap. 3)

PNA: proyecto naturalista amplio (cap. 4)

PSC: psicología de sentido común (cap. 1)

TDA: teoría de la dependencia asimétrica (cap. 3)

TRM: teoría representacional de la mente (cap. 2)

TRCM: teoría representacional/computacional de la mente (cap. 1)



## Introducción

### Representaciones e intencionalidad

Pedro había invitado a Carla y Laura para cenar. *Deseaba* agasajarlas con una comida japonesa y *creía* que a ellas les gustaba el pescado crudo, por eso fue a la pescadería. Pedro tenía la *intención* de comprar un tipo de pescado muy especial pero *temía* que escaseara por aquella época del año. *Decidió* que en caso de que no hubiera el que él *quería*, iba a comprar alguno semejante porque *dudaba* que hubiera tiempo suficiente para preparar otro tipo de comida. Cocinó con esmero un pescado similar al deseado y llegada la hora sonó el timbre de la puerta. Pedro *sabía* que eran ellas y *esperaba* que trajeran el vino especial que habían prometido. Sin embargo, ni bien intercambiaron unas palabras Pedro se decepcionó porque las chicas no habían cumplido su promesa. Ellas *pensaban* comprarlo en un negocio cerca de la casa de Pedro donde siempre tienen ese vino y por eso lo reservaron telefónicamente. Cuando llegaron a la tienda, el vendedor les dijo que hacía media hora lo había vendido porque *creía* que como habían pasado más de dos horas del horario acordado, ya no iban a ir a buscarlo y *deseaba* que otros clientes lo disfrutaran.

Alguna vez, o varias, nos ha ocurrido o hemos escuchado historias como esta. Estos relatos son la manera habitual de explicar las acciones de nuestra vida cotidiana. Lo hacemos tanto cuando damos cuenta de nuestro propio comportamiento como del de los demás. Cuando explicamos las acciones de Pedro, Carla y Laura, en relación con el episodio de la cena, lo que queremos comprender es por ejemplo, ¿por qué Pedro fue a la pescadería?, ¿por qué iba a comprar el pescado crudo que hubiera?, ¿por qué las chicas no trajeron el vino?, ¿por qué el vendedor vendió el vino a otras personas? Nadie espera, ni estaría satisfecho, que ante preguntas como éstas, las respuestas apelarán a los movimientos corporales de los protagonistas de la historia o a sus procesos neurológicos. Si explicáramos,

en nuestras conversaciones cotidianas, por qué Pedro fue a la pescadería esa noche aludiendo a conexiones sinápticas en su cerebro que causaron que su sistema motor enviara ciertos estímulos a las partes del cuerpo involucradas en los movimientos corporales necesarios para llegar a una ubicación espacial en particular, no sólo nos mirarían extrañados sino que, en el fondo, no estaríamos dando respuesta a lo que se nos pide. Porque lo que se nos pide son razones para la acción intencional.

Solemos vernos como seres racionales que actuamos acorde con ciertos propósitos, motivaciones, razones. Lo que nos permite comprender una conducta es hacerla inteligible a la luz de su conformidad con ciertos patrones de racionalidad y para ello apelamos a las razones que la justifican, el “porque” en este caso es, ante todo, conceptual.<sup>1</sup> En líneas generales, una determinada conducta es una acción intencional si hay alguna descripción de esa conducta que apele a términos intencionales, a diferencia de una acción meramente física que se describe en términos puramente físicos (que aluden a respuestas neurofisiológicas y movimientos corporales).<sup>2</sup> A grandes rasgos, el criterio que permite detectar los

---

<sup>1</sup> Aunque por supuesto, se está usando también en sentido causal en tanto que lo que queremos es una ‘explicación’, más allá de una ‘justificación’. Más adelante, también en los capítulos 1 y 2, se verá la cuestión de cómo es posible que las razones sean causas y, en el capítulo 1, cómo podría darse esto sin por ello reducir el ámbito de lo intencional al físico.

<sup>2</sup> Se suelen sumar requisitos más fuertes para la acción intencional como el que la persona se percate de la acción o ésta sea voluntaria (a diferencia de ‘lo que le ocurre’ a la persona) y que la persona esté preparada para dar razones de su acción.

Esta aproximación cruda a la noción de “acción intencional” es neutral con respecto al problema metafísico mente-cuerpo. No dice nada acerca de si una instancia (o tipo) de un evento tiene una descripción física (y en este sentido, es una acción física) y también intencional (y en este sentido, también es una acción intencional) o cada descripción corresponde respectivamente a instancias (o tipos) de eventos mentales y físicos, realizándose los primeros en los segundos, etc. Tal como se verá en los primeros capítulos, los filósofos de la psicología que adhieren a la teoría representacional de la mente suelen adoptar esta última postura metafísica.

Por otro lado, hay problemas conceptuales relativos a la noción de “acción” (por ejemplo, qué es una acción, cómo individuar acciones, cuándo comienza una acción) que pertenecen a la teoría de la acción. Dentro de la filosofía de la psicología se suele adoptar la caracterización general de “acción” como algún cambio en el mundo producido por un agente. Con respecto a la noción de “conducta”, ésta es más abarcativa que la de acción, incluye básicamente (siguiendo a Kim 1996, pp. 28-9): (a) reacciones fisiológicas, (b) movimientos corporales, (c) acciones que involucran movimientos corporales públicamente observables (como ir a la pescadería, comprar vino y demás) y (d) acciones que no involucran movimientos corporales manifiestos (como decidir, calcular, juzgar y demás).

Cuando los filósofos de la psicología hablan de acciones o conductas intencionales (se suele usar “acción” y “conducta” de manera intercambiable y es lo que haré a lo largo de la tesis) y en consecuencia, de explicaciones intencionales, siempre se refieren a (c). El relato de la cena entre los tres amigos es un caso del tipo paradigmático de ejemplos que suelen aparecer en la literatura sobre

términos intencionales paradigmáticos es que son verbos (y sus sustantivos derivados) que requieren de un objeto hacia el cual están dirigidos.

En nuestro relato, los verbos y sustantivos subrayados requieren siempre de un objeto, así, creer requiere que algo sea creído, desear, que algo sea deseado, de manera que no se puede reportar una creencia sin que haya algo creído ni se puede reportar un deseo sin que haya algo que sea deseado. Las oraciones que se utilizan para reportar creencias, deseos, y similares se expresan a través de formas como 'X cree (desea, teme, y similares) que...', los puntos suspensivos se completan con una oración declarativa que expresa una proposición, esto es, la proposición creída o deseada. Dado que los verbos señalan una actitud hacia una proposición, estas expresiones han sido denominadas actitudes proposicionales.

Por ejemplo, la creencia del vendedor de que las chicas no iban a pasar a buscar el vino puede expresarse con la siguiente actitud proposicional: "El vendedor creía que como habían pasado más de dos horas del horario acordado, las chicas ya no iban a ir a buscar el vino", lo que sigue a "que" expresa la proposición creída por el vendedor, el objeto de la actitud de creencia. Se puede expresar el deseo de Pedro de agasajar a sus invitadas por medio de la siguiente actitud proposicional: "Pedro deseaba agasajar a las chicas con una comida japonesa", lo deseado por Pedro es agasajar a las chicas con una comida japonesa (la formulación canónica de las actitudes proposicionales es con el término "que" explícito, aunque aquí no aparece hay maneras de modificar la expresión de modo de que encuadre en la forma canónica).

Hasta aquí he utilizado la denominación "actitud proposicional" para las oraciones que expresan creencias, deseos, y similares. Sin embargo, se suele denominar de igual manera a los *estados mentales* expresados por medio de estas oraciones. Así, los estados mentales como los de creencia, deseo, temor y demás tienen la propiedad de ser intencionales porque siempre son *acerca de* algo. Cuando

---

el tema. Esto no significa que, para los filósofos de la psicología, (d) no sean conductas intencionales y que no deban ser explicadas sino que, para los filósofos naturalistas, mientras que la explicación de (c) es una explicación intencional porque apela a los estados intencionales de (d), la explicación de (d), esto es, de los procesos mentales internos, requiere de explicaciones no-intencionales que den cuenta de cómo son posibles las transiciones causales entre estados mentales (esto se verá en el capítulo 2 cuando se reconstruya la teoría representacional de la mente).

se está en un estado mental de creencia ese estado es acerca de algo, siempre hay algo que se cree; cuando se está en un estado mental de deseo ese estado es acerca de algo, siempre hay algo que se desea, y así con el resto de las actitudes proposicionales. Cuando se habla de estados mentales, la proposición incluida en la cláusula-“que” es denominada el *contenido* del estado mental mientras que se denomina *significado* a la proposición expresada por una expresión lingüística de actitud proposicional (aunque algunos filósofos suelen utilizar ambos términos para hablar del contenido).

Apelar a los contenidos de los estados mentales de actitud proposicional es la manera de dar cuenta cotidianamente de acciones intencionales como preparar una cena, ir a la pescadería, comprar pescado, tocar el timbre, vender vino, y demás. Así, el relato da cuenta de la conducta de los protagonistas en términos intencionales. La explicación intencional consiste en apelar a las creencias, deseos, expectativas y similares de las personas para explicar su conducta: Pedro fue a la pescadería porque “*deseaba* agasajarlas con una comida japonesa y *creía* que a ellas les gustaba el pescado crudo”, compró un pescado similar al que quería porque “*decidió* que en caso de que no hubiera el que él *quería*, iba a comprar alguno semejante porque *dudaba* que hubiera tiempo suficiente para preparar otro tipo de comida”, el vendedor vendió el vino porque “*creía* que como habían pasado más de dos horas del horario acordado, ya no iban a ir a buscarlo y *deseaba* que otros clientes lo disfrutaran”.

Los filósofos de la psicología creen que la psicología cognitiva se ocupa de dar cuenta de la conducta intencional por medio de una versión refinada de este tipo de explicaciones intencionales de sentido común.<sup>3</sup> Ahora bien, para que estas explicaciones de sentido común formen parte legítimamente de una psicología científica se requiere, al menos, (i) cierta concepción realista de los estados de actitud proposicional, (ii) el desarrollo de una teoría acerca de cómo estos estados intervienen en los procesos mentales, y (iii) el desarrollo de una teoría acerca del

---

<sup>3</sup> Cualquiera que esté inmerso en la literatura de la filosofía de la psicología puede advertir que gran parte del esfuerzo está dirigido a sistematizar los enfoques de la psicología cognitiva con los de la psicología de sentido común. El *locus classicus* es Fodor (1975). Pero igualmente esto se puede observar en cualquier texto de esta disciplina (salvo en los textos de los que no son realistas intencionales, en el sentido en que se explicitará más adelante, como Dennett 1987, Stich 1983, P. M. Churchland 1981 y P. S. Churchland 1986).

contenido de las actitudes proposicionales. Según los filósofos de la psicología, la justificación de (i) a (iii) se basa en los desarrollos de la psicología cognitiva en un doble sentido: ésta no sólo supone (i) a (iii) sino que necesita de (i) a (iii) para ser una psicología científica.

Por un lado, las teorías psicológicas, según los filósofos de la psicología, suponen (i) a (iii) porque se puede advertir que son realistas acerca de los estados mentales, sostienen un enfoque de computaciones sobre representaciones y apelan al contenido de los estados intencionales para establecer generalizaciones psicológicas. Por otro lado, requieren (i) a (iii) porque, siempre según los filósofos de la psicología, ser realista intencional es la única manera de que la conducta pueda ser explicada en función del contenido de los estados mentales. Así, tener una teoría representacional computacional de la mente es la mejor explicación disponible de cómo los estados intencionales causan otros estados intencionales y conducta. Y, una teoría del contenido es necesaria para dar cuenta de cómo los estados físicos de un sistema adquieren el contenido al que se apela en las generalizaciones psicológicas para explicar la conducta. Mientras que (i) y (iii) aluden a cuestiones metafísicas (realismo) y semánticas (relación estados mentales-mundo) propias de la filosofía, (ii) parece tener un anclaje más tangible en los desarrollos efectivos de la psicología cognitiva.<sup>4</sup> A continuación veremos brevemente cómo se despliegan (i) a (iii) según la filosofía de la psicología.

Son realistas intencionales quienes sostienen que existen estados intencionales que poseen un contenido que es semánticamente evaluable (es decir, tiene propiedades semánticas tales como condiciones de satisfacción, referencia, significado, y demás) y que causa el comportamiento de los individuos en función de esas propiedades semánticas. La creencia de Pedro de que a sus invitadas les gusta el pescado crudo y su deseo de agasajarlas con una comida japonesa causa que vaya a comprar pescado. Y causa esa acción intencional en función del contenido de sus actitudes proposicionales: el contenido de su creencia son las condiciones en el mundo que harían verdadera la proposición de que a sus

invitadas les gusta el pescado crudo mientras que las condiciones en el mundo que satisfecerían la proposición de agasajarlas con una comida japonesa constituyen el contenido de su deseo. Es porque Pedro tiene la creencia con el contenido que tiene (en este caso, su creencia es verdadera) y tiene el deseo con el contenido que tiene (que en este caso se cumplió) que hizo lo que hizo.

Ahora bien, el problema es cómo hacer que las razones para la acción sean causas de esa acción, siendo que las razones habitan en el reino conceptual de la justificación de la acción mientras que las causas habitan en el reino físico de la explicación de la acción. Los filósofos de la psicología dieron la bienvenida a las teorías de la psicología cognitiva porque la teoría subyacente a las mismas prometía resolver un antiguo problema filosófico que es el de la relación entre el mundo de la causalidad y el mundo de la intencionalidad o cómo lo mental puede causar algo en el mundo siendo que sólo lo físico parece tener poder causal. En su libro fundacional, Fodor (1975) muestra que la teoría que subyace a las teorías psicológicas contemporáneas es la Teoría Representacional de la Mente.

A grandes rasgos, los estados intencionales son estados representacionales instanciados en estados físicos y los procesos mentales son operaciones computacionales/causales sobre representaciones. Los procesos computacionales operan sólo sobre la forma de las representaciones y, en tanto que la forma (o sintaxis) es una propiedad física, éstas causan otras representaciones. Pero, qué operaciones se aplican depende del contenido de las representaciones. Gracias a que la forma codifica el contenido, éste tiene poder causal (una manera muy tosca de entender esto, que se verá en detalle en el capítulo 2, es pensar que las distinciones en contenido reflejan distinciones entre las formas de las representaciones, así dos representaciones con distintos referentes, como PESCADO y VINO tienen distintas formas sintácticas, esto es, las formas de las inscripciones son diferentes).

De manera que, la postulación de un sistema representacional interno parecía solucionar el problema de cómo los estados intencionales que portan contenido

---

<sup>4</sup> Esto no va en contra de que la psicología científica supone y necesita (i) a (iii). Los filósofos de la psicología creen que (i) a (iii) tiene un anclaje en las teorías psicológicas, pero también que (i) y

pueden tener poder causal en virtud de ese contenido, y por ende, permitía que los contenidos pudieran cumplir con ambos requisitos del realismo intencional. Los realistas intencionales, en el sentido explicitado aquí, están comprometidos con la idea de que hay representaciones internas que poseen propiedades intencionales que causan la conducta *qua* propiedades intencionales. De este modo, la filosofía de la psicología vió en la teoría que subyace a las teorías de la psicología cognitiva una salida para dar cuenta del problema de la intencionalidad asimilando sin más la noción de representación con la de intencionalidad. Las representaciones son particulares mentales que son *acerca de* algo en el mundo, de manera que la psicología cognitiva entera se ocupa de explicar la conducta intencional a través de las explicaciones intencionales que apelan a estados intencionales constituidos por representaciones intencionales.

Puesto que de lo que se trata es de dar cuenta de la acción descrita en términos intencionales, la explicación psicológica debe aludir al contenido de los estados intencionales como las creencias y deseos. El tipo de generalizaciones (para algunos, leyes) involucrado sería de la forma: 'Si un individuo desea X, y cree que no puede obtener X, a menos que haga Y entonces, *ceteris paribus*, intentará hacer Y'. La psicología tendría generalizaciones intencionales del estilo "Si un individuo cree que a sus invitados a cenar les gusta el pescado crudo y desea agasajarlos con una comida japonesa, manteniéndose las condiciones iguales, comprará pescado crudo". Así, si son posibles las generalizaciones legaliformes intencionales, entonces la psicología tendrá lo que cualquier disciplina requiere para ser una ciencia autónoma, esto es, leyes especiales de esa disciplina.

Puesto que para dar cuenta del comportamiento humano por medio de explicaciones intencionales, que apelan a generalizaciones legaliformes que relacionan contenidos intencionales, se requiere de la noción de contenido intencional, se concluye que la psicología requiere de una teoría que de cuenta de cómo los estados físicos de un sistema cognitivo adquieren contenido. De manera que el recorrido sería el siguiente: para que haya leyes (o generalizaciones) psicológicas/intencionales, se requiere contenido intencional, éste está conformado

---

(iii) tienen su lugar natural en la filosofía.

por representaciones y para que éstas adquieran contenido se requiere del proyecto de una semántica intencional para las representaciones mentales.

Los estados mentales poseen un contenido y se relacionan con lo representado de la misma manera en que las palabras poseen significados y se relacionan con el mundo. El término "semántica" es trasladado desde la filosofía del lenguaje a la filosofía de la mente para dar cuenta de los contenidos de los estados mentales y de su relación con lo representado. La estrategia supone un aspecto más comprometido que el mero recurso metodológico de la proyección de un término de un ámbito a otro. Es por tener una semántica que las palabras y las oraciones dicen cómo son las cosas y es por tener estados mentales con propiedades semánticas que el sujeto representa cómo es el mundo.<sup>5</sup>

El objetivo de una semántica para los estados intencionales no es el de dar cuenta de por qué una persona tiene las creencias que tiene o cómo se forman las creencias, sino en qué consiste tener un contenido intencional y qué es lo que hace que una creencia sea, por ejemplo, acerca del pescado y no acerca de la carne. La primera cuestión consiste en resolver el problema de cómo encajar las propiedades semánticas/intencionales en un mundo cuyas propiedades son físicas. La segunda cuestión se relaciona con uno de los requisitos que debe cumplir toda semántica del contenido y es el de poder dar cuenta del contenido específico de cada representación.

Si además de ser realista intencional, se es fisicalista, esto es, se cree que las propiedades básicas y últimas son las físicas, entonces el problema a resolver es cómo encajar las propiedades semánticas/intencionales en un mundo cuyas propiedades son físicas. Dicho de otra manera, si desde el punto de vista ontológico se cree en la existencia de estados intencionales con las características mencionadas más arriba y además se cree que la ontología del mundo es física entonces parece ser que se sigue una tesis metodológica que lleva básicamente a naturalizar aquello que en principio no parece ser *per se* parte del mundo natural. Así, las propiedades semánticas

---

<sup>5</sup> En el caso de la semántica informacional de Fodor (que es una de las semánticas de la que me ocuparé) hay una línea directa de explicación que va desde las propiedades semánticas de las entidades lingüísticas a las propiedades semánticas de los estados mentales intencionales y de éstas a las propiedades semánticas de las representaciones mentales.



/intencionales o el contenido de los estados mentales o lo intencional, o lo que sea mental, tiene que poder relacionarse con alguna propiedad del mundo físico.

La intencionalidad, desde esta perspectiva, es una cuestión de vincular representaciones con algo del mundo. De manera que, por un lado, surge la semántica del contenido mental cuyo objetivo es explicar cómo las representaciones internas adquieren significado. Y, por el otro, lo que hay que naturalizar es lo mental que está en el interior del individuo. Como tener un contenido no es una propiedad física y la ontología de los individuos (y del mundo) es física (o la que establezca el conjunto de las ciencias naturales), entonces para que forme parte de la ontología del mundo hay que explicar esta propiedad en términos no-mentales/semánticos/intencionales.

Actualmente, hay dos semánticas naturalistas que pugnan por la mejor explicación del fenómeno de la intencionalidad dentro del marco desarrollado hasta aquí: la semántica informacional y la teleosemántica. Entre las teorías en oferta he elegido las que considero más desarrolladas: la de J. Fodor y R. Millikan, respectivamente. Ambas teorías asumen lo dicho hasta aquí, esto es (i) a (iii). Son teorías que intentan conectar los estados mentales con aspectos del mundo (iii), partiendo de un realismo intencional (i) y una teoría representacional de la mente (ii). Así, los estados mentales son estados internos que poseen un contenido constituido por representaciones que son manipuladas por procesos computacionales/causales en los individuos y que están conectadas con el mundo por medio de propiedades intencionales.

Ambas semánticas son intentos de elaboración de un marco teórico en el cual sea posible ofrecer condiciones suficientes naturales para que un estado físico tenga contenido, esto es, condiciones especificadas en términos no-semánticos y no-intencionales. La idea es reducir (en el sentido de "explicar") la noción de contenido representacional a nociones que apelen sólo a las entidades que las ciencias naturales dicen que hay. Dado que las condiciones suficientes naturalistas se especifican en función del vocabulario de las ciencias naturales, ambas teorías pueden ser calificadas de científicas. Y, en la medida en que comparten una serie de tesis básicas pueden agruparse en lo que denomino el Proyecto Naturalista Científico.

La tesis de doctorado consiste básicamente en una crítica a este proyecto científicista de una semántica naturalista, en la doble vertiente de su enfoque de la intencionalidad y de la psicología cognitiva. La tesis central que intentaré defender es que el proyecto de una semántica intencional para las representaciones mentales es superfluo ya sea que se lo piense para estados del nivel personal como los intencionales o para estados de nivel subpersonal como los de las teorías psicológicas.<sup>6</sup>

Por un lado, si se concibe el fenómeno de la intencionalidad de una manera distinta, relacionado con las prácticas comunicativas y no con la naturalización de la noción de representación, no es necesario el proyecto de una semántica para las representaciones mentales. Si se sostiene que, desde un punto de vista ontológico, un individuo posee estados intencionales (nadie niega que tenemos creencias, deseos, y similares) esto es, se es realista acerca de los estados mentales, pero no se los concibe como una relación con una representación mental que es portadora de propiedades intencionales en el sentido del realismo intencional, entonces no se hace necesaria la tarea de una semántica intencional para el contenido ni mucho menos la de naturalizarlo. Lo que pensamos, creemos, y demás no es una cuestión de qué representaciones tenemos instanciadas en nuestros cerebros y con qué cosas del mundo físico se conectan, sino que son nuestras atribuciones de creencias, deseos, y similares, nuestras expresiones lingüísticas, las que poseen un significado que está conectado con el mundo a través de nuestros intereses, puntos de vistas y, en general, nuestras prácticas comunicativas.

Por otro lado, si se concibe la tarea de la psicología como un intento de dar cuenta de las capacidades subpersonales de los sistemas cognitivos, el tipo de contenido que se requiere no es el intencional y por ende, no se necesita una semántica intencional para las representaciones. En otras palabras, si se cree que el

---

<sup>6</sup> La distinción personal-subpersonal tendrá un papel importante a lo largo de la tesis y se verá en detalle en los capítulos 1 y 5. En términos generales, "subpersonal" alude a aquellos estados que son parte de (sub) sistemas del sistema completo que es la persona, a quien se le atribuyen estados "personales". Por ejemplo, ciertos estados internos del sistema de procesamiento visual, como las imágenes retinales que representan valores de intensidad son subpersonales, mientras que ver una silla no es algo que hace alguna parte de la persona sino que es un estado perceptivo atribuible a la totalidad de la persona.

objetivo de la psicología cognitiva no es dar cuenta de la complejidad de la acción intencional sino de los procesos internos que subyacen al comportamiento, y que el tipo de explicaciones que utiliza no son las intencionales, el proyecto de una semántica intencional para las representaciones de las teorías psicológicas se torna innecesario. Así, intentaré mostrar que la noción de representación que se utiliza, al menos en las teorías psicológicas de las que los filósofos de la psicología se han ocupado con mayor profundidad, no requiere de una semántica intencional.

La idea básica que subyace tanto a la parte crítica como constructiva de la tesis es que debe separarse la cuestión de la representación de la cuestión de la intencionalidad. Los filósofos han cargado a la noción de representación con más peso de lo que ella puede soportar o, en otras palabras, le asignan un rol más ambicioso que aquel para el cual fue pensada. Esta carga atribuida a la noción de representación mental es dar cuenta del fenómeno de la intencionalidad. Así, la filosofía de la psicología al atribuirle a la noción de representación una tarea más pesada y compleja que para la que fue pensada en psicología, motivó la asimilación de la cuestión de las representaciones a la cuestión de la intencionalidad. Pero al efectuar esta asimilación, la filosofía de la psicología concibió de manera inadecuada no sólo la tarea, el alcance y el tipo de explicación de la psicología cognitiva sino el fenómeno mismo de la intencionalidad.

Frente a esta asimilación consensuada en filosofía de la psicología la idea básica es, tal como se dijo más arriba, que hay que separar la cuestión de la representación de la cuestión de la intencionalidad. La psicología cognitiva se ocuparía de dar cuenta de capacidades cognitivas subpersonales apelando a estados representacionales que portan información interna a los subsistemas cognitivos mientras que el fenómeno de la intencionalidad requeriría de un abordaje multidisciplinario que va más allá de los objetivos y alcances de la psicología cognitiva y que se relaciona con las prácticas comunicativas en las que están involucrados nuestro pensamiento y acción. Sin embargo, no es mi intención ofrecer ni una teoría de la representación ni una de la atribución intencional. A lo sumo, la propuesta constructiva se basa en ciertos requisitos para una teoría de la representación y para una de la atribución intencional, y ciertas líneas en favor de que la psicología

cognitiva puede realizar perfectamente su tarea sin la noción de contenido intencional.

A continuación expondré en detalle los contenidos de cada capítulo, pero antes me gustaría sugerir, teniendo en cuenta lo dicho hasta aquí, una línea posible de lectura de la tesis: en el capítulo 1 muestro cómo se hizo la asimilación entre representación e intencionalidad (o exploro (i)). En el capítulo 2 sigo ocupándome de esta asimilación en el marco de la teoría representacional de la mente (o exploro (ii)) y esbozo ciertos requisitos para una teoría representacional en psicología. En el capítulo 3 desarrollo el proyecto naturalista científicista de una semántica para los estados intencionales (o exploro (iii)). En los capítulos siguientes me concentro en (iii). En el 4 discuto el proyecto naturalista científicista y propongo un marco general para dar cuenta de la intencionalidad. En el capítulo 5 exploro la idea de una psicología cognitiva que cumple los requisitos de una teoría de las representaciones sin una semántica intencional. Finalmente, en la conclusión, ofrezco una visión de conjunto a partir de las propuestas de los dos últimos capítulos.

El capítulo 1, tal como se mencionó, está dedicado a mostrar cómo la filosofía de la psicología cognitiva en su afán por encontrar desarrollos empíricos que pudieran ser la base de una psicología de sentido común científica interpretó inadecuadamente las tesis básicas de la psicología cognitiva. El capítulo está dividido en dos partes. En la primera parte, me ocupo del surgimiento de la ciencia cognitiva y de las tesis básicas de la psicología cognitiva. En la segunda parte, me ocupo de la filosofía de la psicología cognitiva y muestro en qué sentidos interpretó erróneamente la tarea de la psicología.

Según una reconstrucción posible, basada en la contraposición con el conductismo y en las fuentes del surgimiento de la ciencia cognitiva, las tesis básicas de la psicología cognitiva consisten, fundamentalmente, en que su objeto de estudio son los procesos internos que subyacen a las capacidades cognitivas, que el nivel de descripción y explicación de estos procesos es en términos de procesamiento de información (mecanismos computacionales y representaciones mentales) y que la explicación del funcionamiento de las capacidades cognitivas consiste en

descomponerlas en subcapacidades más simples. Tal como se mencionó, los filósofos de la psicología vieron en el paradigma cognitivo la posibilidad de incluir el fenómeno de la intencionalidad bajo la ciencia empírica y de esta manera poder solucionar el antiguo problema filosófico de cómo los estados físicos pueden tener la propiedad de ser intencionales. Guiados por este objetivo, desvirtuaron la tarea, el alcance y el tipo de explicación en psicología cognitiva interpretando de manera inadecuada sus tesis básicas. La interpretación de la filosofía de la psicología se equivoca con respecto a la primera y la tercera tesis dado que sostiene que el objetivo de la psicología es dar cuenta de la acción intencional por medio de explicaciones intencionales. Y acierta a medias en la segunda tesis porque si bien describe correctamente los procesos psicológicos tal como se despliegan en las teorías psicológicas, esto es, como computaciones sobre representaciones, sin embargo asimila la cuestión de la representación con la de la intencionalidad.

En el capítulo 2 sigo explorando los ámbitos de la psicología cognitiva y la filosofía de la psicología cognitiva en relación a la asimilación del fenómeno de la representación con el de la intencionalidad y comienzo a delimitar un ámbito representacional no-intencional. En la primera parte, desarrollo y discuto algunas cuestiones básicas que, según los filósofos de la psicología, toda teoría psicológica que postule representaciones tiene que abordar. En la segunda parte, reconstruyo la teoría que la filosofía de la psicología postula como aquella que subyace a todas las teorías de la psicología cognitiva, esto es, la Teoría Representacional de la Mente y evaluo sus respuestas a las cuestiones tratadas en la primera parte.

Las cuestiones básicas acerca de las representaciones mentales se refieren a los criterios para postular representaciones o condiciones para que un sistema sea considerado cognitivo; los requisitos para una teoría de la representación; el formato de las representaciones, esto es, si son imágenes, símbolos, estados neurofisiológicos y/o vectores de activación; la implementación de las representaciones en un sistema físico; y cuestiones acerca de la definición de representación: qué es para algo representar y qué es para algo representar algo específico. En las respuestas de la teoría representacional de la mente a la mayoría de estos temas se ve claramente la asimilación de la cuestión de la representación a la de la intencionalidad.

La respuesta a la primera cuestión sostiene correctamente que el criterio para considerar a un sistema cognitivo es la concepción inferencial de las capacidades cognitivas, sin embargo, las capacidades son caracterizadas intencionalmente. Con respecto a la segunda cuestión, la teoría representacional impone requisitos relacionados con la intencionalidad que no tienen por qué ser necesarios para una teoría de las representaciones. La respuesta a la tercera consiste en el compromiso con un único tipo de formato, que es el simbólico, porque se piensa en los estados intencionales como los estados mentales paradigmáticos. La respuesta a la cuarta cuestión hace que la implementación dependa de la existencia de leyes intencionales, cuestión que se pondrá en duda en el capítulo 4. Por último, los defensores de la teoría representacional sostienen que esta teoría requiere complementarse con una semántica que conecte las representaciones que participan en los estados intencionales con el mundo. Sin embargo, una semántica intencional no es necesaria para una teoría de las representaciones mentales.

Es esta cuestión, la de la semántica intencional, la que ocupará los capítulos restantes. En el capítulo 3 presento el Proyecto Naturalista Cientificista de una semántica intencional. En el capítulo 4 discuto este proyecto con la intención de mostrar que no se requiere una semántica intencional para dar cuenta del fenómeno de la intencionalidad. Y en el capítulo 5 intento mostrar que las representaciones postuladas por las teorías psicológicas no requieren de una semántica intencional. El capítulo 3 está dividido en tres partes. En la tercera, presento las tesis básicas del proyecto naturalista científicista de una semántica del contenido representacional. Las tesis se basan en aspectos centrales de las dos semánticas intencionales contemporáneas más desarrolladas: la Semántica Informativa de J. Fodor de la que me ocupo en la primera parte del capítulo y la Teleosemántica de R. Millikan, de la que me ocupo en la segunda parte del capítulo.

Las tesis básicas del proyecto naturalista científicista abordan tres aspectos: el naturalismo, la naturaleza del contenido, y cuestiones acerca de la psicología de sentido común y la psicología científica. Este proyecto es naturalista en un sentido filosófico y metafilosófico. El naturalismo filosófico con respecto al contenido mental se traduce en la idea de que hay condiciones suficientes para que un estado físico

tenga contenido y que éstas se especifican en un vocabulario no-semántico y no-intencional. El calificativo de científicista agrega a lo anterior la idea de que los términos no-semánticos/intencionales refieren a entidades que las ciencias naturales aceptan en sus teorías. El naturalismo metafilosófico de este proyecto se sigue del naturalismo filosófico y sostiene que la filosofía de la mente se desarrolla en conformidad con y hace uso de la ontología de la ciencia corriente de manera de integrar, desde un punto de vista empírico y conceptual, la mente en el mundo natural científico.

Con respecto al segundo aspecto, el proyecto asume un realismo y un representacionalismo intencional. Tal como hemos visto antes, esto significa que hay estados internos intencionales que tienen un rol causal y son semánticamente evaluables, y cuyo contenido está constituido por representaciones que portan propiedades intencionales. También se asume una determinación unívoca, externalista y atomista del contenido, esto quiere decir que siempre hay un hecho decisivo con respecto a cuál es el contenido de un estado mental, y que éste está completamente determinado por las relaciones del individuo con el mundo, de manera que las creencias asociadas a una creencia en particular no participan de la individuación de esa creencia particular. Con respecto al tercer aspecto, el proyecto sostiene que el punto de partida de la psicología cognitiva es la psicología de sentido común y que uno de los objetivos para el desarrollo de una semántica del contenido mental es lograr una psicología científica.

En la primera parte del capítulo 4 presento cuatro estrategias de ataque a este proyecto: dar argumentos *a priori* con la intención de que sean contundentes y demoledores acerca de por qué un proyecto así no sería viable; formular críticas a las teorías covariacionales y teleológicas, de manera de mostrar que no logran conectar los estados mentales con algo en el mundo y por ende, que el proyecto no ha dado resultados; discutir sus tesis básicas y evaluar cuáles se está dispuesto a aceptar y cuáles no y ver qué tipo de proyecto naturalista quedaría; y oponerle otra concepción de la relación entre el individuo y el mundo, de manera de mostrar que un proyecto de semántica para los contenidos mentales es superfluo. Discuto algunas de las críticas enmarcadas en la primera y la segunda estrategias y opto por desarrollar las

dos últimas estrategias que considero solidarias para proponer una alternativa naturalista: el Proyecto de un Naturalismo Amplio.

En la segunda parte de este capítulo desarrollo las tesis básicas del naturalismo amplio que están estructuradas en torno a los tres aspectos en base a los cuales se presentó el proyecto naturalista científicista. La idea principal es que si se concibe a los estados intencionales de una manera distinta, entonces ya no se requiere de un naturalismo científicista. El naturalismo amplio es realista acerca de los estados intencionales pero no a la manera del realismo intencional. Hay estados internos que portan algún tipo de información que es interpretada en términos semánticos, pero no intencionales, como el contenido de esos estados. Este se determina de manera externalista y holista, y hay casos en los que no hay un hecho decisivo con respecto a cuál es el contenido de un cierto estado mental.

Dado que las representaciones no tienen propiedades intencionales, la intencionalidad no está anclada en el ámbito representacional y por ende no hace falta naturalizar la relación de representación. El fenómeno de la intencionalidad se da en el uso del lenguaje, en el ámbito de la comunicación. Nuestros estados personales intencionales expresados en las expresiones lingüísticas forman parte del mundo natural a través de nuestras prácticas lingüísticas comunicativas que están ancladas en "formas de vida" (Wittgenstein 1953, #241) que se edifican sobre la base de condiciones naturales (sistemas senso-perceptivos en común y regularidades físicas y comportamentales). El naturalismo metafilosófico consiste así en que la filosofía de la mente se desarrolla en conformidad con la ciencia corriente para contribuir a la explicación de los diferentes aspectos del ámbito humano. En cuanto al tercer aspecto, el naturalismo amplio se compromete con la idea de que se puede lograr una psicología científica pero no a partir de la psicología de sentido común.

Una vez que el fenómeno de la intencionalidad queda reubicado en el ámbito de la comunicación, es posible dejar el ámbito de la representación sólo para cuestiones psicológicas. De manera que, en el capítulo 5 intento mostrar que la psicología cognitiva no requiere de la noción de contenido intencional porque ésta se ocupa de mecanismos y estados cognitivos subpersonales no-intencionales. La



noción de contenido apropiada para las teorías psicológicas sería la que denomino 'contenido interno'. Esta noción deflacionista no requiere del desarrollo de una teoría sustantiva, cumple con los requisitos de una teoría de la representación mental tal como se desarrollaron en el capítulo 2, no asimila contenido representacional con contenido intencional y se adecua al objetivo de la psicología de explicar capacidades cognitivas y no la complejidad del comportamiento humano o la acción intencional.

El capítulo está dividido en cuatro partes. En la primera parte, evalúo la distinción doxástico-subdoxástico que se ha propuesto para diferenciar los estados de procesamiento de información y los estados intencionales, y muestro que ésta depende de la distinción personal-subpersonal. En la segunda parte, basándome en aquellas teorías que los filósofos de la psicología consideran paradigmáticas de la psicología (las teorías acerca de las capacidades lingüística y visual cuyos precursores fueron Chomsky y Marr), intento mostrar que el rol explicativo de las representaciones es comprensible a la luz de la idea de que los estados subpersonales poseen un contenido que no es intencional sino interno a los subsistemas de procesamiento de la información. En la tercera parte, intento diferenciar la noción propuesta de contenido interno de la noción de contenido estrecho con la que podría confundirse. Por último, amplío la idea de que son los estados subpersonales aquellos de los cuales la psicología está en condiciones de dar cuenta por sí misma puesto que son estos estados los que intervienen en las explicaciones psicológicas de los mecanismos cognitivos.

La conclusión está destinada a varios fines que abordo a partir de aspectos puntuales. Los fines consisten en ofrecer una visión de conjunto de la propuesta presentada, hacer aclaraciones de algunos de los puntos tratados en los capítulos anteriores, abordar ciertas consecuencias de algunas de las cuestiones planteadas y marcar las diferencias de enfoque con algunas de las nociones y distinciones tradicionales que he utilizado en la tesis. Los aspectos puntuales a partir de los cuales abordaré los fines mencionados consisten, en primer lugar, en la cuestión la asimilación del fenómeno de la representación con el de la intencionalidad. Diré algo más sobre la noción de representación que considero la más adecuada para el ámbito

psicológico, haré algunas observaciones acerca del ámbito de la intencionalidad, y diferenciaré mi propuesta de otras ofrecidas en mi misma línea. En segundo lugar, me ocuparé de la distinción personal-subpersonal. Aclararé en qué sentido la distinción que he utilizado se aparta de la tradicional y me ocuparé de algunas de las consecuencias que se derivan de mi versión de la distinción. En tercer lugar, haré algunos comentarios con respecto a posibles interpretaciones no deseadas de mi tesis principal acerca de la distinción entre el ámbito de la psicología cognitiva computacional y el ámbito de la comunicación. Por último, a partir de lo que se desprende de los capítulos anteriores, intentaré esbozar una respuesta al interrogante del título de la tesis.

*Granny Laura:* no me queda claro si las representaciones mentales son un eslabón entre el individuo y el mundo. Parece que no, pero también parece que sí. De acuerdo a lo que acabás de decir, lo sabremos luego de leer todo lo que prometés en esta introducción, ¿no?

*Autora:* bueno...

*Pedro (filósofo de la psicología):* por supuesto que son un eslabón.

*Auntie Carla (psicóloga cognitiva):* parece que no en el sentido en que vos lo creés.

*Pedro:* y ¿en qué sentido si no?

*Carla:* en el que yo lo creo.

*Pedro:* y ¿cuál es ese sentido?

*Greycat:* a comer todos! que el pescado crudo está pronto!

## 1

## La revolución cognitiva

La noción de representación mental parece tener en la actualidad un protagonismo absoluto y omnipresente, al menos, en aquellos ámbitos relacionados de alguna manera con el estudio de la mente. Esto se debe a que ha tenido un auge expansivo a raíz de la revolución cognitiva que dió lugar, entre las décadas de 1950 y 1970, a la ciencia cognitiva. La ciencia cognitiva es el estudio interdisciplinario de la mente y abarca, fundamentalmente, a la filosofía (en particular, la lógica, la filosofía de la mente y del lenguaje, y la teoría de la decisión), la psicología, la inteligencia artificial, la lingüística y las neurociencias. Junto con la ciencia cognitiva surge un área de la filosofía de la mente: la filosofía de la ciencia cognitiva encargada de abordar cuestiones conceptuales en relación con los fundamentos, la metodología y los problemas de la ciencia cognitiva. La filosofía de la ciencia cognitiva tiene tantas ramificaciones como disciplinas conforman la ciencia cognitiva (filosofía de la inteligencia artificial, filosofía de la psicología cognitiva y demás).

En particular, las disciplinas que nos van a interesar son la psicología cognitiva y la filosofía de la psicología cognitiva porque el objetivo principal de esta tesis apunta a la crítica de un cierto proyecto de la filosofía de la psicología cognitiva que consiste en una semántica para las representaciones mentales que figuran en las teorías de la psicología cognitiva.<sup>1</sup> Este proyecto de la filosofía de la psicología se basa en lo que considero una concepción errónea de los objetivos, alcances, tipo de explicación y concepción de las representaciones de la psicología cognitiva. Dado que la psicología

---

<sup>1</sup> De ahora en más cada vez que utilice "psicología" debe entenderse "psicología cognitiva", lo mismo para la filosofía de la "psicología cognitiva". Por otro lado, se suelen usar como sinónimos las expresiones "filosofía de la psicología" y "filosofía de la mente" aunque, en general, se considera a la filosofía de la psicología como una parte de la filosofía de la mente (aunque véase Block 1981, quien sostiene lo opuesto). Utilizaré estas expresiones indistintamente para referirme a aquel sector de la filosofía que se ocupa de los fundamentos, la metodología y los problemas conceptuales y empíricos de la psicología.

en tanto integrante de la ciencia cognitiva comparte una serie de tesis con el resto de las disciplinas que la conforman, comenzaré por una descripción del marco en el que surgió y se desarrolló la ciencia cognitiva (§1.1 a §1.3). Presentaré sus tesis básicas (§1.4) y haré lo propio con la psicología cognitiva (§1.5) y la filosofía de la psicología (§2.1 y §2.2). Mi objetivo en este capítulo, y en el siguiente, será dejar sentadas ciertas bases que me permitirán desarrollar la idea de que la concepción de la filosofía de la psicología acerca de la tarea de la psicología cognitiva no es la adecuada (§2.3) y que el fenómeno de la intencionalidad no debe asimilarse al de la representación.

## 1. Ciencia Cognitiva

No hay consenso en torno a la fecha específica de nacimiento de la ciencia cognitiva, pero esto es completamente esperable dado que si hay dificultades para hacer lo propio con disciplinas individuales mucho más difícil es establecer el momento en que un conjunto de disciplinas se agrupan en un mismo paradigma. Algunos fechan el nacimiento intelectual de la ciencia cognitiva hacia finales de la década de 1950 (Thagard 2003, Hamlyn 1990), otros alrededor de la mitad de la década de 1960 (Stich 1983) o incluso promediando la década de 1970, si se toma en cuenta su nacimiento organizacional ya que a mitad de esa década surgió la *Cognitive Science Society* en Estados Unidos y la revista *Cognitive Science* (Thagard 2003).

Lo cierto es que la ciencia cognitiva emergió entre las décadas de 1950 y 1970, y es producto de dos fuentes simultáneas. La primera, de psicólogos que reaccionaron contra la psicología conductista imperante en las universidades norteamericanas, la segunda, de fuentes externas a la psicología, a partir de los desarrollos de científicos computacionales que fueron los fundadores de la inteligencia artificial. Esta reseña no pretende ser exhaustiva y en este sentido, dejará de lado muchos aspectos que pueden ser considerados atinentes. Mi objetivo es tan sólo relevar aquellos que servirán para

dar cuenta de las tesis básicas de la psicología cognitiva. Veamos en primer lugar la doctrina oficial contra la que se revelan los psicólogos alrededor de la década de 1950.

### *1.1. El conductismo científico*

El conductismo con figuras como J. Watson, B. Skinner y C. Hull dominó en el período de 1920 a 1950 en los departamentos norteamericanos de psicología. Se puede distinguir entre un conductismo metodológico y uno científico o radical.<sup>2</sup> El conductismo metodológico es una reacción al introspeccionismo y su tesis básica afirma que la investigación psicológica debe basarse en datos conductuales (o, al menos, públicamente observables) vs. los datos de la conciencia (datos privados accesibles por introspección). El conductismo científico radicaliza esta tesis sosteniendo que la tarea de la psicología es la formulación de leyes basadas en relaciones entre estímulos y respuestas observables (vs. cualquier tipo de postulación de mecanismos mentales inferidos). Esta tesis empirista radical se manifiesta en el ámbito de una teoría psicológica sustantiva como la del aprendizaje (área en la que el conductismo enfocó sus investigaciones) en las tesis de que no hay principios innatos que guían el aprendizaje (vs. innatismo) y que éste es producto de mecanismos de propósito general (vs. modularidad).

El conductismo científico es una forma de eliminativismo débil en el sentido de que sostiene que los fenómenos mentales no tienen lugar en una teoría científica.<sup>3</sup> Mientras que el conductismo científico implica el conductismo metodológico, no ocurre lo mismo a la inversa. Sostener que los datos psicológicos deben ser observables públicamente no descarta la postulación de mecanismos (mentales o físicos) inferidos a partir de esos datos. Así, mientras que el conductismo

---

<sup>2</sup> Para ser más precisos habría que distinguir entre el 'conductismo radical' en tanto tesis de la filosofía de la psicología y el 'conductismo científico' en tanto teoría psicológica sustantiva. Así, la tesis básica enunciada más adelante, en sentido estricto, es parte del conductismo radical mientras que más abajo cuando se exponga brevemente las directrices básicas de las teorías psicológicas conductistas me estaré refiriendo al conductismo científico. No me ceñiré a esta distinción estricta porque nada en este capítulo depende de ello.

metodológico es parte integral de la actividad científica de la psicológica cognitiva actual, el conductismo científico fue objeto de ataques demoledores por parte de los incipientes científicos cognitivos (lo que no quita que algunas de sus tesis con respecto al aprendizaje hayan motivado enfoques conexionistas de la arquitectura cognitiva en la inteligencia artificial).

El psicólogo fundador del conductismo científico fue J. Watson que unió en sus últimos escritos el conductismo con la teoría del condicionamiento de I. Pavlov, unión que fundó las bases que dominaron la psicología de la época. Entre los neo-conductistas más influyentes se encuentran B. Skinner y C. Hull. Se puede hablar de dos escuelas que coexistieron, la liderada por Watson y Skinner, y por otro lado, Pavlov y Hull (Gallistel 1998). Los primeros eran radicales en el sentido en que la psicología sólo debía ocuparse de entidades físicas observables. Así, conceptos como memoria, motivación, atención, percepción, no formaban parte de la psicología científica.<sup>4</sup> En cambio, los segundos estaban dispuestos a aceptar ciertas entidades teóricas siempre y cuando se basaran en datos observables y tuvieran encarnadura fisiológica (por ejemplo, la noción de 'fuerza asociativa' basada en la conexión sináptica).

Watson (1913) reacciona frente al introspeccionismo (utilizado como método por E. Titchener y O. Külpe, que fueron alumnos de W. Wundt, fundador de la psicología experimental, y por el funcionalismo americano de W. James) sosteniendo que el estudio de los seres humanos debía enfocarse de la misma manera en que se estudia al resto de los animales, y en ese sentido la psicología pertenece a las ciencias naturales a la par de la física, la química, etc. La tarea de la psicología consiste en el estudio de la conducta observable de manera de establecer leyes y los únicos datos

---

<sup>3</sup> Cf. Skinner (1953), p. 35. Un eliminativismo fuerte sostendría que los fenómenos mentales no existen.

<sup>4</sup> Esto se debe en parte a que Skinner identificaba mentalismo con cartesianismo (que utilizaba "ficciones explicativas", esto es, términos mentales que no tenían referencia) o teorías como la de Freud (cuyas explicaciones mentalistas le quitaban peso tanto al entorno como a la conducta). Skinner en ocasiones adoptaba un conductismo analítico (para el cual los términos mentalistas debían ser analizados en función de disposiciones a la conducta. Cf. Flanagan y Rey 1998) o simplemente un fisicalismo (los términos mentales refieren a eventos cerebrales. Cf. Flanagan 1984, cap. 4).

pertinentes son los públicamente observables.<sup>5</sup> Los últimos trabajos de Watson incorporan los desarrollos de su discípulo K. Lashley sobre condicionamiento (quien conoció los trabajos de I. Pavlov sobre reflejos condicionados) y estos fueron desarrollados a su vez por C. Hull (1940, 1943) en su teoría hipotética-deductiva del aprendizaje denominada Teoría de la Reducción del Impulso. La concepción del aprendizaje es la tradicional del conductismo (la formación de relaciones estímulo-respuesta) pero en su teoría los elementos relevantes en el condicionamiento son las variables del organismo que intervienen entre el estímulo y la respuesta (como impulsos iniciales, entrenamiento pasado, etc.).

Fue Skinner quien mayor influencia tuvo en el movimiento conductista y cuya teoría se ha combatido con mayor encarnizamiento. Primero veremos su teoría psicológica sustantiva y luego su filosofía de la psicología. La Teoría del Condicionamiento Operante de Skinner (1938) se basa en la teoría clásica del condicionamiento de Pavlov (1927) y en la 'ley del efecto' de E. Thorndike. La teoría clásica de los reflejos condicionados intenta dar cuenta de los sistemas mecánicos en términos de cómo funcionan los reflejos incondicionados. Estos responden automáticamente de la misma manera ante los mismos estímulos dado que los reflejos forman parte de nuestro bagaje biológico. Se pueden adquirir nuevas conductas (condicionadas) en función de nuevas conexiones estímulo-respuesta (basadas en las conexiones de los reflejos incondicionados) y a estas conductas se las denomina reflejos condicionados.

El famoso caso de los perros de Pavlov consiste en que se le presenta a un perro, junto con el estímulo incondicionado del olor de la comida, un estímulo condicionado como el sonido de una campana. Después de un período prolongado, la sola presentación del estímulo condicionado produce una respuesta condicionada bastante similar a la respuesta incondicionada de la salivación. Por medio de este tipo de condicionamiento clásico resulta imposible dar cuenta de la conducta novedosa, esto es, cómo dado el mismo estímulo es posible producir respuestas distintas. La novedad sólo se da en el estímulo pero las respuestas son siempre similares (como en el caso de la salivación ante el estímulo nuevo del sonido de la campana). Skinner

---

<sup>5</sup> Watson (1925) constituye el manifiesto conductista de su fundador.

creía que agregando a este tipo de psicología la 'ley del efecto' era posible construir la noción de 'conducta operante' que podía dar cuenta de la novedad comportamental.

La idea fundamental es que se puede dar cuenta de toda la conducta inteligente, tanto en seres humanos como en animales, en función de patrones de condicionamiento que consisten, a grandes rasgos, en que la posibilidad de que ocurra una respuesta  $R$  ante un estímulo  $E$  aumenta o disminuye en función del refuerzo positivo (recompensa) o negativo (castigo) con el que se haya asociado la unidad  $E-R$  en el pasado. La conducta es 'operante' porque se modifica en función de los resultados positivos o negativos del patrón de  $E-R$  (y en este sentido, es conducta regida por el efecto, esto es, por el estímulo que sigue a la respuesta). De esta manera es posible dar cuenta de la emergencia de patrones conductuales novedosos. En última instancia, se puede explicar y predecir la conducta de un organismo en una situación en particular (en términos de probabilidad) en función de su historia ambiental, es decir, en función de los patrones  $E-R$  establecidos a lo largo de la historia del organismo (y en función de las propiedades de su entorno actual, esto es, de la situación en particular).

Skinner (1953) creía que con sólo establecer las leyes físicas que conectaban eventos físicos (los estímulos y las respuestas) se obtenía una explicación de los fenómenos psicológicos sin necesidad de describir lo que ocurre entre el estímulo y la respuesta, esto es, sin necesidad de apelar a procesos internos inobservables. Cualquier teoría que postule entidades inobservables está sujeta al argumento de Skinner que se conoce como el 'dilema del teórico'. Dada una cadena causal respaldada por una ley, cuyo primer miembro es un estímulo  $E$ , el segundo un estado o un proceso mental o neurofisiológico  $M$  y el tercero una respuesta  $R$ , si de lo que se trata es de explicar la conducta observable del organismo, se sigue por transitividad que  $E$  causa  $R$  sin necesidad de apelar a  $M$  (por ejemplo, si hay una ley que conecta la falta de agua en un organismo con el tomar líquido, no es necesario apelar a ningún estado del organismo, mental o físico, que sea 'estar sediento' porque la causa de la conducta es la falta de agua). De manera que, o bien  $M$  está ligado nomológicamente con  $E$  y  $R$ , en cuyo caso, no hace falta referirse a  $M$ , o bien,  $M$  no está ligado nomológicamente con  $E$  y  $R$ , en cuyo caso tampoco hace falta referirse a  $M$ . La



conclusión es que no es lógicamente necesario apelar a procesos internos para dar cuenta de la conducta.

Este argumento por sí mismo no es concluyente a favor de que *E* tenga que participar en la cadena explicativa de *R*, por la misma razón lógica la cadena causal podría ser *M-R*. Para descartar *M* como elemento explicativo de la cadena causal, Skinner apela a los problemas metafísicos/epistemológicos que acarrearía su postulación, tales como si esos procesos internos son inmateriales (como sostenía el dualismo) o son simplemente estados cerebrales (como sostiene el fisicalismo). De manera que apelar a causas internas no sólo era innecesario sino que no hacerlo permitía desembarazarse de problemas metafísicos (y epistemológicos) irresolubles.<sup>6</sup>

### *1.2. Primera fuente: la reacción al conductismo*

Una de las fuentes de la reacción al conductismo surge de psicólogos que ya trabajando en un paradigma distinto comienzan a publicar, alrededor de los mediados de la década de 1950, sus resultados con las consiguientes críticas al paradigma anterior.<sup>7</sup> Los trabajos de G. Miller, D. Broadbent, J. Bruner y N. Chomsky, entre otros, marcan los hitos fundacionales de la ciencia cognitiva. Bruner y Miller fundan en 1960 en Harvard el *Center for Cognitive Studies* y U. Neisser en 1967 publica el primer manual de psicología cognitiva. Influenciados por la teoría de la información de C. Shannon (1948), algunos psicólogos adhirieron a la analogía de la mente como un canal de información. Esta metáfora que le daba un carácter pasivo a la mente fue rápidamente abandonada por una metáfora más

---

<sup>6</sup> Como se dijo en la nota 4, Skinner era fisicalista. El punto es que se oponía a la explicación psicológica reduccionista en términos neurofisiológicos por el problema metafísico mencionado pero además porque consideraba que no sólo el reduccionismo era impracticable sino que aunque fuera posible una teoría de la identidad mente-cerebro de tipos, la descripción de la conducta perdería en riqueza (Cf. Flanagan 1984, cap. 4).

<sup>7</sup> La caída del conductismo no se debe sólo al surgimiento de un enfoque cognitivista en psicología. Hay otros motivos, internos al conductismo, que no expondré puesto que el objetivo es el surgimiento de la ciencia cognitiva y no los motivos de la caída del conductismo.

Por otro lado, hay otras reacciones al conductismo como la psicología Gestalt y la epistemología genética de Piaget de las cuales no me ocuparé (en general, no se las suele considerar como parte de la corriente dominante que constituyó la ciencia cognitiva).

activa y la consecuente analogía de la mente como un canal que procesa información. Al igual que las computadoras, los procesos mentales 'codifican' información (la reúnen y le otorgan un formato representacional), la 'almacenan' y la 'recuperan' cuando es necesario. A continuación enumero algunos de los trabajos que se consideran fundacionales y me concentraré en la crítica chomskiana a la teoría del lenguaje de Skinner.

Miller (1956) propone que las limitaciones en la memoria (los humanos sólo pueden manejar  $7 \pm 2$  ítems de información en la memoria de corto plazo) pueden superarse por medio de representaciones mentales y procesos que operan sobre ellas. También introdujo el concepto fundamental de *chunk* que alude a cualquier unidad significativa de información, que pueden ser dígitos, posiciones en el ajedrez, palabras, etc. Broadbent (1958) propone el primer modelo de los procesos de la atención (Modelo del Filtro). Bruner *et alia* (1956) dan cuenta de la formación de conceptos utilizando la experimentación como indicio de procesos mentales inobservables. Miller *et alia* (1960) discuten el conductismo basados en la analogía mente-computadora con el Modelo de Unidades TOTE (*test-operate-test-exit*), que reemplaza al modelo *E-R* y consiste en testear un objetivo para ver si se lo alcanzó y si no es el caso, se realizan operaciones para alcanzarlo, el ciclo se repite hasta que se alcanza el objetivo o se lo abandona. Este modelo no sólo fue la base de los desarrollos en resolución de problemas, sino que además, incorporó todo un vocabulario mentalista en psicología ("planes", "objetivos", "imagen mental", y demás), de manera que este libro constituyó una especie de manifiesto del nuevo paradigma.

Chomsky (1959) emprende un furioso ataque a la aplicación de Skinner (1957) de sus ideas conductistas al lenguaje, sosteniendo que es imposible dar cuenta de la complejidad de la capacidad lingüística sólo apelando a respuestas conductuales a estímulos físicos.<sup>8</sup> La metodología y el vocabulario de la teoría skinneriana funciona en el laboratorio cuando de lo que se trata es de tareas simples en las que ciertos animales reaccionan con conductas simples (por ejemplo, apretar una palanca para

liberar comida) en respuesta a estímulos sencillos (por ejemplo, el encendido de luces). En estos casos, es posible formular leyes y de hecho han sido el único tipo de leyes que han podido establecer los conductistas. Sin embargo, esta metodología y vocabulario no pueden aplicarse a capacidades complejas humanas fuera de los laboratorios, cuando los estímulos ambientales y las respuestas a los mismos son de una variedad y riqueza extraordinarias.

Según Chomsky, el dilema al que se enfrenta la metodología conductista consiste en o bien se considera 'estímulo' a cualquier evento físico ante el cual el organismo es capaz de reaccionar y 'respuesta' a cualquier parte de la conducta o bien se consideran sólo como estímulos y respuestas genuinas aquellas que se ha demostrado que están respaldadas por leyes. Si se acepta lo primero, eso implica admitir que no hay leyes *E-R* lo cual es negar un postulado fundamental del conductismo. Si se opta por lo segundo, la conducta es legaliforme por definición, con lo cual se dejaría de lado el estudio de la mayor parte de lo que hacen los animales porque no sería conducta al no estar respaldada por leyes. Chomsky analiza en profundidad los conceptos skinnerianos ("estímulo", "respuesta", "control de estímulo", "refuerzo", y demás) mostrando que ni tomándolos en sentido literal ni en sentido metafórico dan lugar a lo que Skinner quería, esto es, una teoría científica acerca del comportamiento lingüístico.

El punto de disputa fundamental en esta polémica es entre la explicación que alude a la historia ambiental del sujeto y la que alude a procesos mentales. Para Skinner una vez que se conoce la historia ambiental del individuo se tiene todo lo que se requiere para la explicación y predicción de la conducta, en cambio, según Chomsky, sólo se puede comprender cómo la historia ambiental afecta a un individuo si se conoce cómo afecta a sus procesos internos. Aplicado esto al lenguaje, Chomsky sostiene que éste no puede consistir en disposiciones a conductas verbales adquiridas por condicionamiento durante la infancia puesto que por un lado, los niños no reciben entrenamiento lingüístico y por el otro, el input es pobre (incompleto, distorsionado, etc.) en relación con la riqueza del output lingüístico (Chomsky

---

<sup>8</sup> En realidad, la intención de Chomsky fue más amplia dado que la crítica apuntaba a toda la psicología conductista. Su objetivo era mostrar, a través de la crítica del libro de Skinner, cómo los

denomina 'problema de Platón' al de la pobreza del estímulo). Además, el uso del lenguaje es creativo en el sentido de ser independiente del estímulo y productivo en el sentido de que es posible comprender y producir un conjunto infinito de oraciones jamás antes escuchadas ('problema de Descartes'). Ambos problemas se explican si se postula un sistema con representaciones y procesos mentales dedicado exclusivamente a la capacidad lingüística.

### *1.3. Segunda fuente: la inteligencia artificial*

La segunda fuente del surgimiento de la ciencia cognitiva se debe al nacimiento de la inteligencia artificial cuyos fundadores fueron J. McCarthy, M. Minsky, A. Newell, y H. Simon, entre otros representantes del cognitivismo clásico, y cuyos trabajos se basaron en los de Turing (1936, 1950). El nombre de la disciplina surge a raíz de un encuentro que organiza McCarthy en 1956, el *Dartmouth Summer Research Project on Artificial Intelligence* en el que A. Newell, J. Shaw y H. Simon presentan su Teórico Lógico, un programa que demostraba teoremas, y que fue el punto de partida para los desarrollos computacionales de la concepción de la inteligencia como la capacidad de resolución de problemas.

La tesis básica de la inteligencia artificial cognitiva ('fuerte') es que una teoría de la inteligencia comprenderá tanto a las personas como a las computadoras pues ambos sistemas son manifestaciones de un mismo fenómeno subyacente. Este fenómeno subyacente es la organización funcional del sistema y se puede dar cuenta de la misma en términos de operaciones computacionales sobre representaciones mentales. Así, el fenómeno general de la cognición depende de la organización funcional y no de qué está hecho el sistema (volveré sobre la inteligencia artificial fuerte y la tesis funcionalista más abajo, y en la segunda parte de este capítulo).

Los dos modelos iniciales de la inteligencia artificial: el cognitivismo clásico y el conexionismo, surgen antes de la instauración oficial de la disciplina, a partir del artículo seminal de McCulloch y Pitts (1947) que integra la lógica proposicional, la

---

supuestos conductistas llevaban a una reducción al absurdo.

teoría de las neuronas de Sherrington y la computación a la Turing. Los modelos computacionales dan lugar a la analogía computadora-mente que postula que la mente es al cerebro lo que un programa es al soporte físico de una computadora, esto es, tanto la mente como los programas se encargan del procesamiento de información con relativa independencia del soporte físico. El cognitivismo clásico postula operaciones computacionales seriales sobre representaciones mentales discretas mientras que el conexionismo se basa en patrones simultáneos de activación computacional sobre nodos de una red, de manera que las representaciones están distribuidas en los nodos.<sup>9</sup>

Mientras que el cognitivismo clásico se desarrolló rápidamente ocupándose de las funciones cognitivas superiores (pensamiento, memoria, juicio, resolución de problemas, y similares), el conexionismo recién en estos últimos veinte años ha emergido con fuerza debido fundamentalmente a las falencias del modelo clásico en relación a las funciones cognitivas inferiores (las perceptivas-motoras). Un libro fundacional del cognitivismo clásico es el de Newell y Simon (1972) en el que se intenta dar cuenta de la resolución de problemas en base a un modelo de sistemas de producción. Un artículo en el que se formula claramente el cognitivismo clásico es el de Newell (1980). Con respecto al conexionismo, en tanto procesamiento distribuido en paralelo, el texto clásico es Rumelhart y McClelland (1986).

#### *1.4. Ciencia cognitiva*

No es una tarea fácil formular las tesis que comparten las disciplinas enmarcadas en la ciencia cognitiva porque, en primer lugar, sigue siendo un campo de desarrollo continuo con lo cual los límites que podrían marcar las tesis están sujetos a revisión constante ante la posible incorporación de disciplinas secundarias a las cinco principales que la conforman. Y, en segundo lugar, parece haber cierta tensión entre algunas de ellas (por ejemplo, las neurociencias con respecto al resto de

---

<sup>9</sup> En los modelos conexionistas no hay programas en el sentido tradicional del cognitivismo clásico pero la distinción básica a la que apunta la metáfora (información vs. soporte físico) se

las disciplinas, esto se verá más abajo). Intentaré formularlas en contraposición a lo que hemos visto de la corriente conductista y basándome en las fuentes del surgimiento de la ciencia cognitiva. Las tesis principales del cognitivismo son las siguientes:

*1. Tesis acerca del objeto de estudio:*

Las ciencias cognitivas son ciencias de la mente y en tanto tales se ocupan de dar cuenta de las capacidades cognitivas como la memoria, el lenguaje, la percepción, etc. Su objetivo es explicar los mecanismos internos (los estados y procesos mentales) que subyacen a (o causan) la conducta.

En contra del conductismo que consideraba la psicología como la ciencia de la conducta, para el cognitivismo el objeto de estudio no es la conducta en sí sino aquellos procesos que median entre los estímulos y las respuestas. Las respuestas (aquello que había que explicar para el conductismo) sirven como datos de los cuales extraer inferencias acerca de los procesos mentales (lo que hay que explicar para el cognitivismo).

*2. Tesis acerca de la concepción de la mente.*

La mente es un sistema de procesamiento de la información y en tanto tal, el nivel adecuado para su descripción es en estos términos.

La mente es un sistema que codifica, almacena, transforma y recupera información. La noción de información viene de la Teoría de la Información de Shannon (1948) cuyos conceptos fundamentales son los de emisor, receptor y un canal intermediario por el cual se trasmite el mensaje. El mensaje puede codificarse y cuantificarse en unidades discretas que portan información (bits). La noción de información consiste, a grandes rasgos, en aquello que modifica el estado de incertidumbre del receptor. Con el desarrollo de la inteligencia artificial surge la idea de 'sistemas formales automáticos interpretados' (Haugeland 1985), esto es, las

---

mantiene igual.

computadoras, que consisten básicamente en la manipulación algorítmica de información.

La tesis 2 puede subdividirse en cuatro tesis:

2.1. El nivel de procesamiento de la información es un nivel intermedio entre las descripciones de sentido común y las que apelan al sustrato físico.

El cognitivismo postula que hay un nivel de descripción de lo que hace un sistema que es, justamente, el nivel cognitivo en términos de procesamiento de la información. Este es entendido como operaciones computacionales sobre representaciones mentales y es un nivel distinto del nivel de las descripciones del sentido común o fenomenológicas y las que apelan al sustrato físico del sistema. Las descripciones del sentido común son las que utilizamos cotidianamente para la comprensión y predicción de nuestra conducta y la de los demás y se formulan apelando a términos mentalistas. Cuando las utilizamos no se supone que nos comprometemos con ninguna afirmación acerca de los mecanismos internos que subyacen a la conducta. Por otro lado, las descripciones en términos del sustrato físico de un sistema si bien apelan a mecanismos internos son formuladas en términos físicos. El nivel intermedio de descripción de la función de un sistema cognitivo apela a mecanismos internos y en su formulación figuran términos mentalistas.

Ahora bien, hay otros sistemas aparte de los humanos que pueden ser descritos como procesadores de información, como las computadoras y algunos animales. En este sentido, se puede entender la analogía con la computadora de dos formas. En la versión débil (tesis 2.2 de más abajo), la analogía provee un nivel de descripción, cierto vocabulario y una manera de poner a prueba ciertas teorías psicológicas por medio de la simulación de las capacidades cognitivas. En la versión fuerte (tesis 2.3 junto con la tesis 2.4 de más abajo), es posible elaborar una teoría cognitiva general que abarque tanto a los humanos como a las computadoras. Esto sería posible porque la cognición depende de la organización funcional del sistema y si se puede lograr en una computadora la misma estructura funcional que poseen los

humanos entonces éstas tendrían una mente al igual que la de los humanos. De esta manera habría un nivel, el del procesamiento de la información, en el que ambos sistemas tendrían la misma estructura funcional. Estos sentidos de la analogía pueden formularse de la siguiente manera:

2.2. La analogía con la computadora provee el vocabulario y el nivel de explicación adecuado para una teoría cognitiva.

2.3. El nivel de procesamiento de la información es común a todos los sistemas que pueden ser descritos en términos computacionales como procesadores de información (humanos, computadoras y ciertos animales).

La tesis funcionalista 2.4. afirma lo siguiente:

2.4. Las capacidades cognitivas son caracterizadas funcionalmente de manera independiente de su realización material.

Lo que le interesa a la ciencia cognitiva es la organización funcional de la mente con abstracción de su implementación física. Esto no quiere decir que no forme parte de la ciencia cognitiva una teoría de la implementación de las capacidades cognitivas o que la evidencia proveniente de las neurociencias no sea pertinente (ver la tesis 3.1 más abajo). De hecho, las neurociencias forman parte de la ciencia cognitiva y los datos que aportan son fundamentales a la hora de dar cuenta del funcionamiento de las capacidades cognitivas. Lo que se quiere resaltar es que se pueden desarrollar teorías funcionales de manera independiente y como un primer paso para las teorías de implementación. Así, mientras haya una buena descripción de lo que hace un sistema cognitivo será posible indagar qué elementos físicos del sistema llevan a cabo esa función.

Es la tesis funcionalista la que adosada a 2.3 permite sostener que reproduciendo la estructura funcional de un sistema, todas las propiedades mentales tendrían que reproducirse. Así, una máquina, tal como se supone que es el cerebro,



con la estructura funcional adecuada tendría que poder reproducir pensamientos, creencias, en fin, todo lo que nosotros con nuestros cerebros pensamos y sentimos. Puesto que el funcionalismo es la teoría mente-cerebro de la filosofía de la psicología que subyace a esta tesis de la ciencia cognitiva, me ocuparé de ella en detalle en la segunda parte del capítulo.

La última tesis que comparten las disciplinas que participan de la ciencia cognitiva es la 3 que tiene dos aspectos, 3.1 y 3.2 :

### 3. Tesis acerca de la explicación cognitiva:

3.1. Cualquier teoría cognitiva requiere de, al menos, tres niveles de explicación: el computacional, el algorítmico y el implementacional.

Marr (1982, 1977) distingue tres niveles de explicación desde los cuales entender la/s tarea/s que lleva a cabo cualquier sistema de procesamiento de información. (1) En el *nivel computacional* se ofrece un análisis abstracto (formal) del problema que va a ser computado. Se relaciona con la pregunta de qué va a ser computado y por qué. El análisis se dirige a la naturaleza del problema para el cual la teoría es una solución y ésta es entendida a la manera de una teoría de la competencia. (2) En el *nivel algorítmico* se especifica la representación para el input y el output, y el algoritmo para la transformación. Y, (3) en el *nivel implementacional* se describe cómo puede realizarse físicamente la representación y el algoritmo.

Según Marr, los niveles no son absolutamente independientes entre sí desde un punto de vista lógico y causal. El nivel (1) tiene cierta independencia o neutralidad en relación al nivel (2). En cambio, el nivel (2) depende del (3). Y, en algún sentido, tanto (2) como (3) imponen ciertas restricciones al nivel (1) (por ejemplo, en términos del límite de información que se puede almacenar o si el sistema funciona en serie o en paralelo). La tarea de una teoría cognitiva consiste en caracterizar a un sistema de manera de entender cómo puede lograr X donde X refiere a cualquier logro cognitivo humano. En relación a esta caracterización del sistema, algunos investigadores consideran que el objetivo principal de una teoría

psicológica consiste en la especificación de la competencia del sistema. Es decir, una formulación abstracta de la función cognitiva que computa.

En este sentido, algunos teóricos defienden, en términos de Sterelny (1990, p. 198), una 'autonomía metodológica' que sostiene que no hay que preocuparse por la implementación computacional o física, pues, en principio, no aportaría nada sustancial a la descripción de la capacidad del sistema. O bien, otros defienden una 'autonomía epistémica' (Sterelny 1990, pp. 198-9) sosteniendo que las teorías psicológicas sólo requieren de evidencia psicológica. Aunque esto no significa afirmar que la evidencia de otras ciencias es irrelevante, es una manera más de afirmar que los detalles relacionados con la implementación no enriquecen la teoría psicológica.

Sin embargo, cualquier teoría que esté bajo el cognitivismo asume la tesis 2, esto significa que los modelos, al menos, se implementan computacionalmente. Si bien puede ocurrir que la teoría misma no esté en condiciones de proporcionar una teoría de la implementación computacional o física, deberá adherir a alguna de manera de mostrar la viabilidad de la descripción del nivel computacional. En el capítulo siguiente, en el que me ocupó de las representaciones mentales, volveré sobre el tema de la implementación ya que este aspecto tendrá un papel importante en relación con los requisitos que cualquier teoría acerca de las representaciones mentales debe satisfacer.

Veamos ahora el segundo aspecto de la explicación cognitiva:

3.2 La explicación en el nivel computacional consiste típicamente en un análisis funcional.

No me extenderé sobre este tema ya que en el capítulo 4 expondré en detalle en qué consiste el análisis funcional. Sólo a manera de introducción diré que el análisis funcional (Cummins 1983, Botterill y Carruthers 1999, intr.) consiste en la descomposición de las capacidades complejas en subcapacidades simples que explican la capacidad analizada. Este análisis suele reflejarse en los diagramas de flujo

o de cajas en los que estas últimas representan las subcapacidades, y las flechas que van entre las cajas marcan el flujo de la información o el algoritmo pertinente.<sup>10</sup>

Hay otras tesis que podrían agregarse a estas tres pero que, si bien algunas son ampliamente aceptadas, otras no tanto y por eso es preferible no incluirlas dentro de las tesis básicas del cognitivismo. Son tesis opcionales en relación a la arquitectura mental. Toda teoría cognitiva debe encuadrarse, en última instancia, bajo alguna de las arquitecturas mentales que actualmente están disponibles. Lo que sostenga la teoría sustantiva dependerá de cómo se conciba el modelo de procesamiento de la información, si como lo hace el cognitivismo clásico, el conexionismo u otras arquitecturas en oferta.

Por otro lado, muchas de las arquitecturas mentales postulan que las capacidades cognitivas son modulares (vs. que son mecanismos de propósito general como sostenían los conductistas) adscribiendo así a alguna de las versiones de la tesis de la modularidad (Fodor 1983, Pinker 1994, Chomsky 1995b). La teoría sustantiva acerca de cierta capacidad debe decir algo acerca de si esa capacidad es modular o no pero, por supuesto, no tiene por qué comprometerse con que el resto de las capacidades sean también modulares (modularidad masiva) o con la idea de que hay módulos y sistemas centrales (modularidad parcial). Otro aspecto discutible es la tesis innatista. El cognitivismo nace con ciertos compromisos fuertes con la idea de que existe información innata en el sujeto (vs. el conductismo), y algunos modelos de las capacidades cognitivas adscriben a la misma (por ejemplo, Marr 1982). Sin embargo, no la considero una tesis central ya que no es una condición necesaria para que una teoría esté bajo el paradigma cognitivista.

Una última cuestión con respecto a la denominación habitual de cognitivismo para los desarrollos de la ciencia cognitiva. Este término tiene un doble sentido. El cognitivismo en sentido amplio, agrupa a disciplinas que estudian

---

<sup>10</sup> El análisis funcional no mantiene ninguna relación con la orientación funcionalista en psicología (fundada en la década de 1890 en la Universidad de Chicago) que se centra en un enfoque evolucionista de adaptación de las funciones mentales al entorno. El análisis funcional es un tipo de explicación para las capacidades cognitivas instanciadas computacionalmente (véase la nota 28 para la forma que adopta el análisis funcional) y surge bajo el marco del funcionalismo filosófico (que se verá en la segunda parte de este capítulo), el cual es neutral con respecto a los enfoques teleológicos de las capacidades cognitivas.

los mecanismos internos que subyacen a las capacidades cognitivas y éstas se entienden como cualquier función de la que es capaz la mente. Esto acuerda con el espíritu bajo el que surgió la ciencia cognitiva y que quedó plasmado en, por ejemplo, la definición de "cognición" del primer manual de psicología cognitiva: "cognición" refiere a todos los procesos por los cuales el input sensorial es transformado, reducido, elaborado, almacenado, recobrado y usado." (Neisser 1967, p. 4)

El sentido estricto alude a la idea del estudio de la cognición como un ámbito distinto de la percepción y las emociones. Esta acepción se debió en parte a que los primeros logros de la ciencia cognitiva (en particular los obtenidos bajo el cognitivismo clásico en inteligencia artificial) enfocaron sus investigaciones en las funciones cognitivas superiores como el pensamiento y la resolución de problemas dejando de lado las funciones cognitivas inferiores. En las últimas décadas se ha vuelto a la concepción amplia de la fundación de la ciencia cognitiva con el estudio de la percepción (en particular, la visión) y las emociones. Es en este sentido amplio que se entenderá el cognitivismo. Veamos a continuación algunas cuestiones específicas en torno a la psicología cognitiva.

### *1.5. Psicología cognitiva*

La psicología cognitiva es la disciplina central en el paradigma cognitivista. Es por ello que no sólo comparte las tesis principales de la ciencia cognitiva sino que se podría decir que le otorga sus tesis básicas a la ciencia cognitiva. De modo que las tesis básicas de la psicología cognitiva son las que hemos visto en la sección anterior y que pueden resumirse de la siguiente manera:

#### *1. Tesis acerca del objeto de estudio.*

La psicología cognitiva se ocupa de dar cuenta de los procesos mentales que subyacen a las capacidades cognitivas.

2. *Tesis acerca de la concepción de la mente:*

El nivel adecuado de descripción de los procesos mentales es en términos de procesamiento de la información.

3. *Tesis acerca de la explicación psicológica:*

Las explicaciones típicas de los procesos mentales en términos de procesamiento de la información tienen la forma del análisis funcional.<sup>11</sup>

La psicología cognitiva se ocupa de dar cuenta de los procesos mentales que participan en capacidades cognitivas como (formuladas de la manera en la que suelen aparecer en los manuales) percepción, memoria, atención, lenguaje, representación del conocimiento, resolución de problemas, toma de decisiones, aprendizaje, pensamiento y emociones. A pesar de que en sentido amplio, la psicología abarca cualquier desarrollo cognitivista, como la psicología social cognitivista, la psicología del desarrollo cognitivista, y la psicolingüística, las relaciones entre la primera y estas disciplinas no siempre son de interacción (Garnham 1994).

No es mi intención marcar límites al ámbito de la psicología cognitiva producto de disputas internas metodológicas o de lo que la corriente dominante considera que son las teorías que se enmarcan legítimamente bajo esa denominación. Al contrario, creo que toda disciplina que satisface las tesis enunciadas más arriba tiene pleno derecho a formar parte de la misma. A continuación, simplemente mostraré una de las disputas internas porque me

---

<sup>11</sup> Tal como presento la cuestión, se podría pensar que no habría diferencias entre la psicología y cualquier otra disciplina de la ciencia cognitiva dado que toda disciplina que satisfaga las tres tesis básicas se convertiría automáticamente en psicología. Por supuesto que no es esto lo que quiero decir. Las tesis de la psicología que a su vez son la base para considerar a cualquier ciencia como cognitiva son condiciones necesarias, cada ciencia a su vez (incluso la psicología misma) tendrá tesis que constituyan condiciones suficientes para ser considerada tal ciencia en particular. Me interesan sólo las tesis básicas de la psicología (y no las particulares) porque considero que son esas las que la filosofía de la psicología interpretó de manera inadecuada.

interesa en particular, para los desarrollos de la psicolingüística, dejar sentado el criterio de satisfacción mencionado más arriba.<sup>12</sup>

Con respecto a las relaciones entre la psicolingüística y la psicología cognitiva, algunos psicólogos cognitivos consideran que la primera está fuertemente influenciada por los desarrollos en lingüística (en particular, la chomskiana) mientras que las teorías del procesamiento agrupadas en la 'psicología del lenguaje' no le deben nada a esta disciplina teórica y en este sentido, algunos psicólogos suelen no considerar a la psicolingüística como parte de la psicología cognitiva.<sup>13</sup> No acuerdo con esta divisoria de aguas. Por un lado, esta discrepancia es una disputa puramente interna que no socava las bases del cognitivismo, esto es, el estudio de los mecanismos y representaciones internos que subyacen a las capacidades cognitivas (más el resto de las tesis mencionadas más arriba). Y el modelo chomskiano se inscribe, sin duda, bajo esta directriz.

Por otro lado, aquellos que acordamos en que los desarrollos de la lingüística chomskiana que intentan dar cuenta del funcionamiento de un mecanismo en particular, la facultad del lenguaje, son también modelos de procesamiento (que deben ser incluidos en modelos más generales, esto es, que lleguen hasta el output final que consiste en la producción o comprensión de las emisiones lingüísticas), creemos que es completamente legítimo su estatus de modelo de la psicología cognitiva. De manera que, no sólo el modelo chomskiano está pensado para dar cuenta de los mecanismos internos del lenguaje sino que considero que es un modelo paradigmático para establecer el alcance, el tipo de explicación y los objetivos de la psicología cognitiva. Esto se verá en detalle en el capítulo 5.

---

<sup>12</sup> Con respecto a la psicología social, hay ciertas diferencias de criterios metodológicos que hacen que no se considere a ésta como parte de la psicología cognitiva a pesar de que estudia aquellos aspectos cognitivos del comportamiento social. Por ejemplo, la psicología cognitiva le crítica a la psicología social cognitiva su metodología poco rigurosa y ésta última le critica a la primera el aislamiento en los laboratorios. Con respecto a la psicología del desarrollo si bien hay mayor interacción con la psicología cognitiva, la primera está más influenciada por la filosofía de la mente (sobre todo en teorías de la mente en los niños) que por la psicología cognitiva misma.

<sup>13</sup> En particular, se le discute a la lingüística chomskiana la postulación de un módulo específico del lenguaje y el innatismo (véase Reber 1987 como ejemplo de - no sorprendentemente- argumentos pobres en contra de la psicolingüística). No voy a entrar en esta discusión, en parte, por las razones expuestas a continuación.

Dado que la filosofía de la psicología ha reflexionado en particular (por no decir casi por completo) sobre la teoría de la visión cuyo precursor es D. Marr y la teoría psicolingüística cuyo precursor es N. Chomsky, y dado que las tomaré como teorías paradigmáticas de la psicología cognitiva para establecer parte del punto principal que quiero defender,<sup>14</sup> las voy a exponer de manera resumida. Esto también permitirá tener una idea acerca del tipo de teorías sustantivas de la psicología cognitiva que se ajustan perfectamente a las tesis básicas de la misma. Comenzaré con la teoría chomskiana.

### *1.5.1. Dos teorías de la psicología cognitiva*

Al mismo tiempo que Chomsky emprende su crítica al conductismo, en (1957) desafía la lingüística estructuralista americana (L. Bloomfield, Z. Harris) que dominaba por aquella época y concebía la tarea de la lingüística como el descubrimiento de procedimientos mecánicos que aplicados a un corpus daban como resultado una gramática descriptiva. La lingüística generativa se enfrenta a la lingüística taxonómica estructuralista concibiendo la tarea de la lingüística como el estudio de los mecanismos y representaciones mentales que subyacen a la capacidad lingüística. Se postula un módulo específico en la mente encargado del procesamiento de la información lingüística que es denominado facultad del lenguaje. Esta tiene dos componentes: un sistema cognitivo y sistemas de ejecución (sensorio-articulatorios e intencionales-conceptuales).

Se deja a un lado el estudio de los sistemas de ejecución que constituyen el uso del lenguaje porque son de una complejidad extrema, en los que intervienen múltiples factores que no pueden ser aislados y simplificados para su estudio y en

---

<sup>14</sup> Tal como se mencionó en la introducción, lo que quiero mostrar es que no hace falta una semántica intencional para las representaciones mentales que figuran en las teorías psicológicas (capítulo 5). La otra parte de la defensa se basa en que no se requiere de una semántica intencional para dar cuenta de la intencionalidad (capítulo 4).

A decir verdad, en el capítulo 5 me ocuparé en profundidad de la teoría chomskiana y haré sólo algunas comparaciones con la teoría marriana para las cuales, sin embargo, se requiere de cierta familiaridad con la teoría. Es por ello que la presento brevemente aquí.

esa medida son poco proclives para una investigación científica (Chomsky 1991, p. 40). De manera que, la lingüística se concentra en el estudio del sistema cognitivo, de su estado inicial (la gramática universal conformada por un conglomerado de rasgos fonológicos, semánticos y sintácticos parametrizados que se postulan que son innatos) y su estado estable (lenguaje-I). El lenguaje-I tiene como inputs instancias de eventos físicos (ondas sonoras) y sus outputs son objetos abstractos (objetos simbólicos o formales), que llevan en sí toda la información semántica, sintáctica y fonológica necesaria para asociar el sonido y el significado de las expresiones lingüísticas. Los componentes del lenguaje-I que permiten transformar el input en output son un mecanismo computacional y un léxico (Chomsky 1995a, 1995b).

Cada lenguaje-I consiste en un procedimiento computacional (inferencial) que genera descripciones estructurales de las expresiones lingüísticas a partir de elecciones léxicas. El procedimiento computacional toma ítems léxicos cuyos rasgos fonológicos se computan para dar lugar a una representación fonológica (o Forma Fonológica (FF)), toma rasgos semánticos para dar lugar a una representación semántica (o Forma Lógica (FL)) y los rasgos formales entran en la computación pero deben ser eliminados en el camino a la FF para dar lugar a una representación convergente (interpretable por los sistemas de ejecución). De manera que, una expresión lingüística E (el output del lenguaje-I) consiste, al menos, en un par {FON, SEM} formado por una derivación convergente óptima. FON(E) es una representación de la FF que contiene información relevante al sonido de E y SEM(E) es una representación de la FL que porta información relevante al significado de E. Estas representaciones actúan como instrucciones para los sistemas de ejecución que son los sistemas de interface que imponen condiciones de legibilidad puesto que son los que interpretan el output del lenguaje-I y siguen sus instrucciones para el pensamiento y la acción.

La teoría de la visión de Marr (1982) tiene como objeto la descripción de cómo el sistema visual construye una representación tridimensional de objetos a partir de las matrices grises que son las imágenes retinales. Así, el mecanismo visual tiene como inputs pares de imágenes retinales que representan valores de



intensidad y como outputs representaciones de objetos de tres dimensiones. El procesamiento de la información se realiza en tres etapas que conforman tres niveles de representación. La primera se denomina 'esbozo primario' y consiste en la construcción por procedimientos computacionales a partir de patrones de luz y sombra de una representación de cambios de reflectancia. Esta representación corresponde al nivel de la retina (que es una superficie de dos dimensiones) y la información alude a lo que sucede en este nivel. A partir de esta representación el sistema construye el 'esbozo 2.5D' que representa información acerca de la orientación y profundidad de las superficies y sus contornos en un marco de coordenadas centrado en el observador. Por medio de transformaciones sobre esta representación, el sistema construye la 'representación de modelo 3D' que describe las formas y organización espacial en un marco de coordenadas centrado en el objeto.

Marr considera que el pasaje de una etapa a la otra es inferencial, pero la información explícita que porta la representación que sirve como premisa no es suficiente para la etapa posterior de manera que se postula que el sistema tiene supuestos *a priori* (innatos) que le permiten realizar las inferencias. Estos supuestos innatos no tienen que estar representados explícitamente en el sistema (esto es, estarían 'cableados' en los mecanismos perceptivos). Un ejemplo de este tipo de supuesto es el siguiente (pp. 214-5): Marr se pregunta "¿cuáles son los supuestos que es razonable adoptar y que empleamos de modo inconciente cuando interpretamos siluetas...como formas tridimensionales?" y uno de los supuestos es que "cada línea de mirada desde el observador hasta el objeto debe incidir en la superficie de éste exactamente en un punto". De manera que cada representación se construye a partir del input del nivel inferior, ciertos supuestos y un algoritmo computacional que de lugar al output que constituye esa representación.<sup>15</sup>

Como se puede notar en estas brevísimas exposiciones, ambas teorías se inscriben bajo las tesis básicas de la psicología cognitiva dado que 1. estudian los procesos internos que subyacen a las capacidades lingüísticas y visuales, 2. lo hacen

---

<sup>15</sup> Para buenas presentaciones de la teoría de Marr, véase Sterelny (1990), cap. 4; Hamlyn (1990), cap. 4, y Segal (1989).

en términos de procesamiento de información (mecanismos computacionales y representaciones mentales) y 3. explican el funcionamiento de las capacidades descomponiendo la capacidad lingüística y visual en capacidades más simples. Con respecto a esto último, ambas teorías descomponen la capacidad de la que se ocupan y enfocan su estudio en un aspecto o subsistema de la misma. Mientras que la teoría chomskiana se ocupa del subsistema del language-I cuyo output es el input para otros subsistemas de la capacidad lingüística global, la teoría marriana se ocupa de un aspecto del sistema perceptivo de la visión, que es el espacial, cuyos outputs son los inputs de otros sistemas que darían lugar a la percepción global.

Además, ambas teorías son un intento de resolver el típico problema de la pobreza del estímulo para el cual la postulación de procesos y representaciones mentales internos es la solución que distingue a las teorías cognitivas (volveré sobre esto en el capítulo siguiente). Para ello, se apela a tres tesis que suelen figurar en las teorías cognitivas (aunque las primeras dos no siempre figuran, tal como hemos visto más arriba): la modularidad, el innatismo y la inferencialidad. Ambas consideran que las capacidades que estudian son modulares en el sentido básico de la especificidad de dominio (el sistema procesa sólo un determinado tipo de inputs), postulan ciertos mecanismos e información innata para que las inferencias sean posibles y conciben a las capacidades como capacidades inferenciales (me ocuparé de esto último en el capítulo siguiente).

A continuación (y también en el capítulo siguiente) veremos cómo la filosofía de la psicología en su interpretación de la psicología transforma alguna de las tres tesis principales de la psicología cognitiva dando lugar a las confusiones que intentaré clarificar en los capítulos finales.

## 2. Filosofía de la Psicología Cognitiva

La revolución cognitiva fue recibida con mucho entusiasmo por parte de un sector de los filósofos de la mente que vieron en el nuevo paradigma no sólo la posibilidad de la resolución del problema de la intencionalidad sino la posibilidad de lograr una psicología científica que reivindicase la psicología de sentido común. Para abordar ambas cuestiones es necesario ocuparnos del fenómeno de la intencionalidad.

### *2.1. Intencionalidad*

Brentano (1874) reintroduce la concepción aristotélica-tomista de la 'inexistencia intencional' de un objeto en la mente sosteniendo que los fenómenos psíquicos se caracterizan por la 'referencia a un contenido', 'la dirección hacia un objeto' o la 'objetividad inmanente'. Esto quiere decir que todo fenómeno psíquico contiene un objeto, por ejemplo, en el juicio hay algo afirmado o negado, en el amor hay algo amado, etc., este objeto intencional es el contenido del estado psíquico. Según Brentano, todos los datos de nuestra conciencia se dividen en físicos y psíquicos. Los fenómenos físicos son las experiencias sensoriales producto de la percepción externa de objetos o modificaciones de una parte del cuerpo, como el dolor. Los fenómenos psíquicos son ideas o dependen de ellas y su taxonomía consiste en ideas (presentan algo a la mente), juicios (se afirma la existencia o no de algo que se presenta a la mente) e intereses (actitud a favor o en contra que acompaña a los juicios, por ejemplo, emociones, sentimientos, deseos, voliciones).

La propiedad de contener un objeto intencional es exclusiva de los fenómenos psíquicos y en tanto tal es lo que los diferencia de los fenómenos físicos, al punto tal que es el criterio para su taxonomía de lo mental dado que ésta depende de cómo el individuo está relacionado intencionalmente con el objeto de sus estados psíquicos. Brentano concebía la intencionalidad como algo parecido a una relación entre el individuo que está en el estado mental, y el contenido del acto mental. No es una relación estricta porque para que lo sea ambos *relata* tienen que existir, pero con respecto a las relaciones mentales puede darse el caso de que el

objeto acerca del que se piensa no exista.<sup>16</sup> Dado que los fenómenos psíquicos conforman una clase que no es idéntica a la de los fenómenos físicos puesto que poseen la propiedad de la intencionalidad, los primeros son irreducibles a los segundos. Y, en la medida en que lo mental es irreducible (afirmación que se conoce como la 'tesis de Brentano'), la psicología es una ciencia autónoma en un sentido metafísico.<sup>17</sup>

La psicología, según Brentano, se divide en 'descriptiva' y 'experimental'. La descriptiva, que luego se denominó 'fenomenología', es en la cual se enmarca toda su doctrina acerca de la intencionalidad. Es una disciplina *a priori* que se ocupa de la formulación de leyes universales acerca de la mente atendiendo a los elementos de la conciencia y cómo se combinan. La experimental es empírica y se ocupa de formular leyes causales acerca de lo mental y de sus relaciones con lo fisiológico. La psicología descriptiva (o fenomenología) no requiere de datos experimentales ya que la fuente de la evidencia de la estructura de la conciencia es la percepción interna (autonomía epistémica), ni requiere acudir a detalles de implementación de la psicología experimental (autonomía metodológica).

Algunos de los análisis de Brentano en torno al fenómeno de la intencionalidad han sido desarrollados, en particular, por las corrientes fenomenológica (Husserl) y analítica. Dentro de esta última, todos los filósofos de la mente han aceptado la idea fundamental de Brentano de que la intencionalidad es la marca de lo mental, en el sentido de que no puede abordarse el ámbito de lo mental sin abordar el fenómeno de la intencionalidad. La mayoría de estos filósofos consideran que la intencionalidad es una relación entre un individuo y el contenido de su estado mental,<sup>18</sup> y sostienen algún sentido de autonomía de la psicología.

Así, la intencionalidad es la propiedad relacional de los estados mentales de ser *acerca de* algo. Los estados mentales paradigmáticos son los de creencia y

---

<sup>16</sup> Brentano modificó su concepción de la intencionalidad, primero rechazando la idea de que el objeto es inmanente para aceptar que el objeto puede estar fuera de la conciencia, y luego rechazando la posibilidad de objetos no existentes. En la actualidad, se considera que un estado intencional es acerca de algo del mundo exterior.

<sup>17</sup> La 'autonomía metafísica' es el tercer sentido de autonomía (junto con la epistémica y la metodológica que vimos antes) que distingue Sterelny (1990), cap. 9.

<sup>18</sup> Muy pocos han sostenido que es una propiedad monádica. Véase el capítulo siguiente.

deseo, pero también hay otros estados mentales que son intencionales como los de pensamiento, conocimiento, intención, duda, espera, y similares. Por ejemplo, estar en un estado de creencia es estar relacionado con el contenido de esa creencia. La relación se denomina actitud, de manera que se puede estar relacionado con un mismo contenido a través de distintas actitudes (por ejemplo, creer y desear que llueva) o estar relacionado con distintos contenidos a través de la misma actitud (por ejemplo, creer que llueva y creer que el pasto es verde). El contenido es la proposición que figura en la cláusula-que que sigue a la actitud (por ejemplo, Alex cree que el pasto es verde). Por eso se denomina a los estados intencionales 'Actitudes Proposicionales' (de ahora en más AP) (Russell 1940).

A diferencia de Brentano, hoy en día la taxonomía de lo mental distingue entre estados intencionales y estados cualitativos. Estos últimos son estados en los que se experimentan ciertas cualidades particulares de las experiencias concientes y es lo que los filósofos llaman 'lo que se siente' al estar en ese estado (esto es, al tener un dolor, al ver algo, etc.). Hubo (y sigue habiendo) una gran polémica, de la cual no podré ocuparme aquí, en torno a si se puede hablar de contenido en los estados cualitativos. Sin embargo, parece haber cierto consenso en que esto es posible de manera que la tesis central de la intencionalidad como la marca de lo mental parece seguir en pie.<sup>19</sup> No obstante, en lo que sigue se hablará de estados intencionales aludiendo a las APs puesto que son los estados intencionales paradigmáticos.

Son realistas intencionales quienes sostienen que existen estados mentales en los organismos con un contenido que es semánticamente evaluable (es decir, tiene propiedades semánticas tales como condiciones de verdad, referencia, significado, y demás) y que causa el comportamiento de los individuos en función de esas propiedades semánticas. Que sean semánticamente interpretables quiere decir que poseen propiedades semánticas, en particular, condiciones de satisfacción: el contenido de las creencias tiene condiciones de verdad y el de los deseos tiene condiciones de cumplimiento. Por ejemplo, el contenido de la creencia

---

<sup>19</sup> Algunos de los que sostienen que hay un aspecto intencional en los estados cualitativos son Tye (1996), Dretske (1995) y Lycan (1996).

de Alex de que el pasto es verde será verdadero o falso según que el pasto sea verde o no. Y el contenido del deseo de Alex de tomar una cerveza se cumplirá o no según si Alex toma la cerveza o no. Las propiedades semánticas permiten hacer generalizaciones puesto que estas últimas se definen sobre elementos que tienen valores de verdad. Así, las generalizaciones que involucran contenidos figuran en el *explanans* de las explicaciones de la conducta de los individuos. De esta manera, se puede explicar (o predecir) la conducta de los individuos de comprar sillas verdes para su jardín porque 'creen que el pasto es verde' y 'desean que sus sillas tengan el mismo color que el pasto'.

Esto se conecta con la idea de que los estados intencionales poseen un papel causal en la producción de la conducta en el sentido de que estos estados son condición suficiente para producir un cierto comportamiento. Así como explicamos la rotura de un frasco de vidrio con agua cuando se lo calienta apelando al factor causal del momento del impacto de la colisión de las moléculas del agua con la superficie del frasco, podemos explicar la conducta de Alex al tomar una cerveza en el bar apelando al factor causal de la conjunción de su deseo de tomar una cerveza y su creencia de que hay cerveza en el bar.

Incluso, se dice que no necesitamos ser filósofos para suscribir a estas propiedades de los estados intencionales porque están implícitas en las explicaciones de sentido común del comportamiento de los individuos que utilizamos constantemente en nuestros intercambios cotidianos con las personas. Estas explicaciones dan cuenta del comportamiento en función del rol causal del contenido de los estados intencionales. Así, se dice que estos intercambios suponen una Psicología de Sentido Común (PSC, lo que se suele llamar *folk psychology*) que usamos para explicar y predecir nuestro propio comportamiento y el de los demás.

Algunos filósofos consideran que sólo un sistema representacional interno puede cumplir con ambos requisitos del realismo intencional. Sólo los elementos que poseen propiedades semánticas (como valores de verdad) podrán hacer las veces del contenido de los estados intencionales y sólo los elementos que poseen propiedades físicas podrán en tanto vehículos del contenido tener un rol causal en la conducta. Dado que se concibe que todo contenido es intencional en tanto posee valores de

verdad, se suele asimilar las propiedades semánticas a las propiedades intencionales, y por ello se habla de propiedades semánticas/intencionales.<sup>20</sup> En cuanto a las propiedades físicas, una propiedad que puede tener efectos causales es la sintaxis, y ésta es entendida como una propiedad abstracta de la forma de ciertos elementos (Fodor 1987, p. 40).

Así, el sistema representacional postulado por la psicología cognitiva viene a cumplir el papel de aquellos elementos con propiedades semánticas y sintácticas que el defensor de la PSC necesita para que las entidades postuladas por la PSC puedan ser incorporadas a la psicología científica. Dicho de otra manera, si la psicología cognitiva postula representaciones que poseen propiedades semánticas y sintácticas, esto permitiría (junto con otras tesis que veremos más adelante) incorporar las explicaciones de sentido común a la psicología científica dado que las propiedades que se les atribuye a los estados intencionales desde la PSC serían las mismas que aparecerían en los estados postulados por la psicología científica. Más adelante veremos cómo en la Teoría Representacional/Computacional de la Mente (TRCM) se conectan ambos tipos de propiedades de manera que las semánticas/intencionales puedan tener un rol causal pero, en principio, los realistas intencionales, en el sentido explicitado aquí, están comprometidos con la idea de que hay representaciones internas que poseen propiedades semánticas/intencionales que causan la conducta.<sup>21</sup>

Tal como Brentano planteó la distinción entre los fenómenos mentales y los físicos parece haber creado un hiato entre las ciencias encargadas de dar cuenta de los mismos. Hemos visto que, por un lado, estarían las ciencias naturales (incluida la psicología 'experimental') y por el otro la psicología 'descriptiva'. Esta última gana en

---

<sup>20</sup> Esta asimilación es explícitamente adoptada por todos los filósofos de la psicología. Hasta donde sé, el único que no asimila propiedades semánticas a las intencionales es Cummins (1989), quien distingue entre contenido semántico y contenido intencional, siendo el primero el contenido de las representaciones entendidas en tanto estructuras de datos y el segundo el contenido que hay que agregar a las representaciones cuando figuran en estados intencionales (también en Cummins 1996 se distingue contenido representacional o semántico y contenido intencional en los estados intencionales). Volveré sobre esto en los capítulos finales dado que es uno de los puntos clave para el objetivo de la distinción entre el ámbito representacional y el intencional.

<sup>21</sup> Aunque no todos los realistas acerca de los estados intencionales se comprometen con esta caracterización (tanto de los estados postulados por la PSC como de los estados postulados por la psicología), por eso "realista intencional" se utilizará de manera estricta tal como se definió (por ejemplo, Dennett (1991a) se considera a sí mismo, de manera correcta según mi parecer, realista

autonomía y en el estatus de científica pero a costa de un dualismo indeseado. Desde la perspectiva brentaniana, el fenómeno de la intencionalidad queda relegado a una investigación *a priori* sin que parezca tratable desde un punto de vista empírico. Los filósofos de la psicología vieron en el paradigma cognitivo la posibilidad de incluir el fenómeno de la intencionalidad bajo la ciencia empírica y de esta manera poder solucionar el problema de cómo los estados físicos pueden tener la propiedad de ser intencionales. Para abordar esto veremos las tesis principales del funcionalismo, un enfoque de la filosofía de la mente que da cuenta de la naturaleza de la mente.

## 2.2. *Funcionalismo filosófico*

El funcionalismo es una teoría metafísica acerca de la naturaleza de la mente que da una respuesta a la pregunta: ¿qué son los estados mentales? La respuesta breve es que los estados mentales son estados funcionales que se realizan en estados físicos. Se puede distinguir entre un funcionalismo metafísico-epistemológico y uno metafísico-ontológico.<sup>22</sup> El funcionalismo epistemológico sostiene que los estados mentales se caracterizan en términos de los roles funcionales que poseen. La postura ontológica que suele adoptar el funcionalismo es un fisicalismo de casos sin reducción. Veamos primero en qué consiste el funcionalismo ontológico.

Los estados mentales son estados internos que se caracterizan por el rol funcional que poseen en la economía cognitiva de un organismo. El rol funcional de un estado mental se especifica en función de las conexiones que ese estado tiene con

---

porque cree que existen estados intencionales pero no poseen un rol causal *qua* estados intencionales).

<sup>22</sup> El funcionalismo es una teoría metafísica a la par del dualismo, el conductismo filosófico y la teoría de la identidad. Distingo un aspecto ontológico porque si bien la respuesta metafísica a qué son los estados mentales difiere de las otras posturas mencionadas, el funcionalismo acuerda con las dos últimas en una ontología fisicalista. Por otro lado, distingo un aspecto epistemológico porque, tal como veremos, dentro del funcionalismo hay diferencias con respecto a la individuación de los roles funcionales. En lo que sigue hablaré simplemente de funcionalismo epistemológico y ontológico (esta distinción respeta en sustancia la de Flanagan 1984, p. 216, quien denomina epistemológico a lo que denominé metafísico-epistemológico y metafísico a lo que denominé metafísico-ontológico).



aquello del entorno que lo causa, con otros estados de los cuales es su causa o efecto y con la conducta causada por ese estado. La caracterización funcional se da en un nivel abstracto de descripción de una función. Muchas cosas, aparte de los estados mentales, tienen descripciones funcionales. Por ejemplo, una silla sirve para sentarse no importa si está hecha de madera o hierro (por supuesto que si está hecha de papel no satisfecerá esta descripción funcional). Así, distintos tipos de cosas según los distintos casos pueden satisfacer el mismo conjunto de especificaciones que determinan el rol funcional de un estado mental.

El funcionalismo fisicalista sostiene que los estados mentales se realizan en un sustrato físico, pero no importa en qué tipo de sustrato siempre y cuando éste satisfaga el rol funcional atribuido.<sup>23</sup> Así, algo es una mente siempre y cuando posea una cierta organización funcional que se realiza en sistemas físicos ya sea que estos estén compuestos, por ejemplo, de neuronas o siliconas. Esta idea es lo que se conoce como la 'tesis de la realizabilidad múltiple' (Putnam 1967) que sostiene que un estado mental-tipo puede realizarse no sólo en sistemas físicos diferentes (personas o robots) sino en distintos sistemas hechos de lo mismo (humanos o animales, o dos humanos) e incluso en diferentes estados físicos de un mismo organismo (un mismo ser humano). Por ejemplo, el estado mental-tipo que consiste en la creencia que hace frío puede realizarse en un ser humano y en un robot o en Julia y Diana, en distintos estados neurofisiológicos-tipo, e incluso se puede realizar en distintos estados neurofisiológicos-tipo de Diana en las distintas ocasiones en que tiene esa creencia. De manera que no se pueden establecer identidades entre los estados mentales-tipo y los estados físicos-tipo.<sup>24</sup>

---

<sup>23</sup> El funcionalismo es compatible con posturas no-fisicalistas, pero hasta donde se sabe nadie defendió esta conjunción, esto es comprensible dado que el problema para el que se propone el funcionalismo es el de cómo los estados mentales se conectan con los estados cerebrales.

Entiendo el fisicalismo como la tesis que sostiene, al menos, que todas las propiedades de las entidades (objetos, eventos, personas, etc.) son o están determinadas por propiedades físicas y en ese sentido, son entidades físicas. El funcionalismo fisicalista se compromete entonces con la idea de que los estados mentales (o mecanismos) internos que figuran en la especificación de los roles funcionales tienen una base física.

Por otro lado, la noción de realización no ha sido objeto de desarrollos detallados pero basta aquí con entender que un evento (o propiedad) A se realiza (o instancia o implementa) en otro evento B cuando B es el sustrato (o correlato) físico de A.

<sup>24</sup> El funcionalismo es una reacción a la teoría de la identidad de tipos cuyos precursores fueron Smart (1959) y Armstrong (1968). Esta última sostiene que lo que hace que un estado mental sea,

En general, los funcionalistas se comprometen con una teoría de la identidad de casos que establece que una instancia de un estado mental-tipo es idéntica con una instancia de un estado físico-tipo, pero no hay identidad de tipos, básicamente por la tesis de la realizabilidad múltiple.<sup>25</sup> Esta imposibilidad de la identidad de tipos hace del funcionalismo una teoría fisicalista no-reduccionista con lo que no sólo respeta la tesis de Brentano sino que permite sostener una autonomía metafísica de la psicología. Los filósofos creen que en tanto que los fundamentos teóricos de la psicología son funcionalistas, es decir, en tanto que la psicología se basa en una caracterización funcional de los estados mentales, es posible sostener que las clases naturales de la psicología no son idénticas a las clases naturales de las ciencias naturales de manera que el descubrimiento de los ocupantes de los roles funcionales no tornará descartable a la psicología. En el caso de sistemas físicos con propiedades funcionales, como las mentales, se requiere tanto de la descripción física como de la funcional, esta última especifica los roles funcionales y la primera da cuenta de cómo estos se realizan.

El funcionalismo epistemológico sostiene que los roles funcionales son roles causales y estos se especifican en función de las relaciones causales que mantienen los estados mentales entre sí, con inputs perceptivos y outputs motores.<sup>26</sup> Así, por

---

por ejemplo 'dolor' (lo que tienen en común todos los estados de dolor) es tener una cierta estructura física, mientras que el funcionalismo sostiene que es tener un cierto rol funcional. En términos de propiedades, la diferencia radica en afirmar que las propiedades mentales son propiedades físicas de primer orden *vs.* que son propiedades de segundo orden (porque se cuantifica sobre propiedades de primer orden. Así, una propiedad mental consiste en tener una propiedad con un cierto rol funcional).

<sup>25</sup> En términos de propiedades, la teoría de la identidad de tipos sostiene que toda propiedad mental es una propiedad física y la de casos que cualquier evento que tiene una propiedad mental también tiene una física. Para una formulación de las diferencias entre ambas, véase Kim (1996), cap. 6.

<sup>26</sup> Todos los funcionalistas epistemológicos acuerdan en que los roles funcionales ya sea que se los considere roles conceptuales, inferenciales, o computacionales tienen que tener una base o mecanismo causal subyacente, por eso la formulación del funcionalismo epistemológico es en términos de relaciones causales. Sin embargo, hay diferencias entre estos tipos de roles que no podré abordar aquí, véase Cummins (1989), cap. 9.

Por otro lado, se puede distinguir entre un funcionalismo amplio (externalista) y estrecho (internalista) en función de la individuación de los inputs y los outputs. El primero sostiene que se individualizan en función del entorno o las causas distales (un ejemplo de input sería: 'ver un vaso con agua' y de output: 'agarrar el vaso con agua'). El estrecho sostiene que se caracterizan en función de las causas y efectos proximales (un ejemplo de input sería la descripción del estímulo retinal y de output, la descripción de la orden motora a los músculos de los ojos y las manos). Volveré sobre la distinción tradicional entre externalismo e internalismo en el capítulo 3.

ejemplo, se puede individuar la creencia de que llueve porque el organismo que la posee está en algún estado neurofisiológico que está causado típicamente por la visión de (el escuchar, etc.) la lluvia y que causa típicamente junto con otros estados mentales (por ejemplo, el deseo de salir de casa y el deseo de no mojarse) la conducta de salir con un paraguas.<sup>27</sup> Las distintas versiones del funcionalismo difieren en cómo se especifican esas relaciones causales. Se puede distinguir dos enfoques que agrupan las distintas versiones: los que describen los roles funcionales a partir de los conceptos de la PSC y los que lo hacen a partir de los conceptos de la psicología científica.

El funcionalismo de la PSC (Lewis 1972, Armstrong 1968) considera que en primer lugar es necesario un análisis conceptual (o funcional) de los términos mentalistas con el objetivo de elucidar sus significados, y en una segunda etapa empírica se examinan qué objetos pueden ser los ocupantes de los roles causales analizados conceptualmente en la primera etapa.<sup>28</sup> La ventaja de especificar el rol causal a partir de la PSC es que sus conceptos y generalizaciones parecen ser sumamente estables en comparación con los de las teorías psicológicas, pero si lo que se quiere es una teoría científica acerca de los estados mentales, sería más apropiado ir más allá de los análisis de sentido común de manera que los mismos puedan ser contrastados empíricamente.<sup>29</sup>

---

<sup>27</sup> Este ejemplo supone una individuación externalista de sentido común (ver más adelante y nota anterior).

<sup>28</sup> No confundir el análisis funcional del funcionalismo de sentido común para dar cuenta de la conducta intencional con el análisis funcional en tanto explicación en psicología cognitiva de las capacidades cognitivas. Si bien ambos tienen una etapa de análisis en el sentido de que hay descomposición en elementos más simples, lo que se descompone en el primero son los términos mentales definidos según ciertas afirmaciones triviales de la psicología del sentido común mediante la técnica Ramsey-Lewis con el objetivo de formular definiciones funcionales en las que no figuren expresiones psicológicas (la forma de este análisis es: "Cuando alguien está en tal y tal combinación de estados mentales y recibe estimulación sensorial de tal y tal clase, él tiende con tal y tal probabilidad a ser causado mediante ello a pasar a tales y tales estados mentales y a producir tales y tales procesos motrices" (Lewis 1972, p. 212). Lo que se descompone en el segundo son capacidades cognitivas complejas en función de los elementos más simples, todos postulados por las teorías psicológicas científicas (la forma de este análisis sería algo así: la propiedad P del sistema cognitivo S tiene el análisis [C1... Cn, O], donde C son los componentes y O la manera en la que están organizados. Véase Cummins (1983), p. 17 y el capítulo 4 de esta tesis.

<sup>29</sup> El punto es que la corrección de la especificación de los roles funcionales depende de que la teoría, cuyos conceptos intervienen en las especificaciones, sea verdadera. Muchos han cuestionado no sólo que la PSC sea una teoría (Dennett 1987) sino la verdad de sus generalizaciones (Churchland

De las versiones que no parten de un análisis de sentido común de los roles de los estados mentales, el funcionalismo de tabla de máquina (Putnam 1960) sostiene que los estados mentales cumplen el rol de ser estados de una máquina de Turing. Poseer una mente es tener implementado en el cerebro una máquina de Turing, de manera que la relación entre el soporte físico de la máquina y el programa es análoga a la de los estados cerebrales y los mentales. Esta versión ha recibido numerosas críticas que dieron lugar a un enfoque más refinado que llamaré funcionalismo computacional.<sup>30</sup> Este abarca tanto el enfoque clásico como el homuncular.

El funcionalismo homuncular (Lycan 1981, Dennett 1978) se propone como una manera de dar cuenta de la intencionalidad descomponiendo las capacidades cognitivas (/intencionales) en subsistemas cada vez más tontos que serían los homúnculos.<sup>31</sup> Así, las capacidades cognitivas (/intencionales) se descompondrían cada vez más hasta llegar a los niveles no-intencionales, de manera que la intencionalidad quedaría explicada por estados y mecanismos no-intencionales. En general, los defensores de este enfoque agregan una perspectiva teleológica cuando adoptan un enfoque evolucionista acerca de la intencionalidad y sostienen que los estados intencionales son rasgos adaptativos del organismo diseñados para cumplir funciones.

El funcionalismo computacional clásico (Pylyshyn 1980, 1984; Fodor y Pylyshyn 1988) sostiene que las relaciones causales entre los estados mentales consisten en operaciones computacionales sobre los constituyentes de estos estados, es decir, las representaciones. Esta versión del funcionalismo está reflejada en la TRCM (en filosofía de la psicología) o, lo que es lo mismo, en el 'Cognitismo Clásico' (en inteligencia artificial), y se propone como una manera de entender lo que he caracterizado como la tesis 2. Me ocuparé en detalle de la TRCM en el capítulo siguiente dado que no sólo constituyó uno de los fundamentos de la ciencia cognitiva

---

1981). Una objeción menos radical a este enfoque es la de ser demasiado liberal en la atribución de estados intencionales, ver Block (1978).

<sup>30</sup> Las críticas aluden a que hay características de los estados mentales que este tipo de funcionalismo no recoge. Por ejemplo, no es posible diferenciar estados ocurrentes de los disposicionales, la individuación de estados se torna demasiado fina, y los estados de los humanos son infinitos a diferencia de los de una máquina de Turing, véase Block (1978) y Fodor (1980).

<sup>31</sup> La denominación 'homuncular' proviene de una descripción caricaturesca de las funciones cognitivas como realizaciones llevadas a cabo por hombrecillos que poseemos en la cabeza.

sino que aún, a pesar de las múltiples críticas recibidas, sigue siendo el paradigma dominante en la psicología cognitiva.<sup>32</sup> Aquí, simplemente haré una presentación general centrada en cómo la TRCM vió en la fundación de la psicología cognitiva la posibilidad de solucionar los problemas que mencioné al comienzo de este apartado y cómo guiada por esos objetivos interpretó de manera errónea ciertas tesis básicas de la psicología cognitiva. La razón principal de esta interpretación errónea radica en la asimilación del problema de la representación al de la intencionalidad. Para mostrar esto es preciso adentrarnos en cuestiones que requieren de cierto desarrollo, por eso en este capítulo dejaré planteada la cuestión y me ocuparé del desarrollo detallado de los argumentos en los capítulos restantes.

La TRCM sostiene que los roles causales de los estados mentales son roles computacionales. Así, un estado mental de creencia se diferencia de uno de deseo porque está sujeto a distintos procesos computacionales. Los procesos computacionales consisten en secuencias causales de transformaciones (algoritmos) sobre símbolos en virtud de las propiedades estructurales (formales, sintácticas) de los mismos. De manera que los constituyentes de los estados mentales son símbolos que tienen propiedades sintácticas. Pero además, tienen propiedades semánticas dado que constituyen el contenido de los estados mentales. Puesto que las explicaciones intencionales aluden al poder causal de los contenidos para dar cuenta de la conducta de los individuos, concebir el contenido como símbolos instanciados en el cerebro (en esto consistirían las representaciones mentales) permite sostener que es gracias a las propiedades sintácticas de las representaciones que éstas tienen poder causal. Y dado que las propiedades sintácticas están correlacionadas con las propiedades semánticas, en última instancia, son las propiedades semánticas (el

---

<sup>32</sup> Esto no significa que no haya teorías psicológicas contemporáneas desarrolladas bajo otros modelos de la arquitectura mental. Por ejemplo, un grupo de críticas específicas al funcionalismo clásico viene de otros enfoques contemporáneos, provenientes de la inteligencia artificial, como el conexionismo (véase la polémica entre los defensores del conexionismo y el cognitivismo clásico reimpresa en la parte I de MacDonald y MacDonald 1995) y la teoría de sistemas dinámicos (para una presentación filosófica véase van Gelder 1995). El punto es que dado que la ciencia cognitiva surge con el enfoque de la arquitectura clásica y éste sigue siendo aún el paradigma dominante, los filósofos de la psicología han intentado conciliar sus propuestas de semántica con este enfoque. Puesto que me centraré en la cuestión de la semántica para las representaciones mentales, conviene tener presente la concepción de la mente del cognitivismo clásico.

contenido de los estados mentales) las que tienen poder causal en la producción de la conducta.<sup>33</sup>

Así, la TRCM resuelve los problemas mencionados al comienzo de este apartado. Permite que el fenómeno de la intencionalidad sea tratable en el marco de una teoría científica porque los estados intencionales son estados funcionales que se instancian en estados físicos, de manera que la intencionalidad no sólo es posible en un mundo de objetos físicos sino que es posible lograr una ciencia empírica de lo mental. Y dada la tesis de la realizabilidad múltiple es posible sostener la irreducibilidad de lo mental, y con ello la autonomía de la psicología científica. Además, la TRCM une la PSC con la idea de computación a la Turing, de modo que lo que se computan son los contenidos de los estados mentales que tienen las mismas propiedades que la PSC les atribuye (tienen poder causal y son semánticamente evaluables). De manera que las explicaciones intencionales de sentido común pueden transformarse en explicaciones de la psicología científica. Así, la idea de computaciones sobre representaciones mentales no sólo provee la base metafísica de la psicología científica sino también la metodológica.<sup>34</sup>

---

<sup>33</sup> Para lo que me interesa remarcar a continuación no hace falta desarrollar cómo las propiedades sintácticas *reflejan* las semánticas, eso se verá en el capítulo siguiente. Por ahora basta con la idea de que las computaciones sobre la sintaxis de los símbolos respetan ciertas propiedades semánticas que los mismos mantienen entre sí, como por ejemplo la que se da entre las premisas y la conclusión de un razonamiento válido.

<sup>34</sup> Quisiera aclarar que el funcionalismo clásico y el homuncular, si bien difieren en algunos aspectos, son perfectamente compatibles. Ambos son expresiones del intento de dar cuenta de la intencionalidad en términos de procesos mecánicos computacionales sobre representaciones, de manera que para ambos el cerebro es una máquina sintáctica que produce semántica (entendida esta última como intencionalidad). El enfoque jerárquico (desde el nivel intencional hacia el nivel neurofisiológico) postulado por el funcionalismo homuncular es compatible con la perspectiva de que cada sistema y subsistema computa sobre representaciones mentales hasta el nivel anterior al neurofisiológico. El aspecto teleológico generalmente agregado al funcionalismo homuncular es completamente independiente de este último, de hecho, este aspecto puede agregarse a cualquier concepción de la mente. En cuanto al compromiso del funcionalismo clásico con representaciones mentales que son símbolos con propiedades sintácticas y semánticas instanciados en el cerebro, un funcionalista homuncular podría adherir a esta concepción, las razones para no adherir a la misma son completamente independientes del funcionalismo homuncular en sí. Lo mismo con respecto a las razones para sostener que el funcionalismo computacional tiene que incorporar la PSC a la psicología científica. Se puede sostener un cognitivismo clásico sin adherir a la idea de que este enfoque es una descripción del mecanismo interno que subyace a las atribuciones de la PSC.

Por otro lado, las críticas al funcionalismo computacional pueden agruparse en aquellas que apuntan a que el funcionalismo no explica el fenómeno de la intencionalidad, en particular, la objeción de la habitación china (Searle 1980) y la objeción de chauvinismo (Block 1978), y aquellas que apuntan a que no explica el fenómeno de la conciencia, en particular la objeción de los *qualia* (Nagel 1974, Jackson 1982, Block 1978).

### *2.3. Funcionalismo computacional y psicología cognitiva*

Hemos visto que las tesis básicas de la psicología cognitiva consisten en: 1. [objeto de estudio] estudiar los procesos internos que subyacen a las capacidades cognitivas, 2. [concepción de la mente] hacerlo en términos de procesamiento de información, y 3. [explicación psicológica] explicar el funcionamiento de las capacidades descomponiéndolas en subcapacidades más simples. De la sola comparación entre las tesis de la psicología cognitiva y esta propuesta de la filosofía de la psicología surge que la TRCM respeta la tesis 2 a medias porque asimila estados representaciones a estados intencionales e interpreta erróneamente 1 y 3 porque cree que el objetivo de la psicología es explicar la acción intencional y que las explicaciones intencionales son las explicaciones típicas en psicología cognitiva. Veamos una por una las tesis.

La tesis 2 afirma que es posible describir la mente en términos funcionalistas como un sistema cuya función es procesar información. Para dar cuenta del procesamiento de la información se postula un nivel de descripción de la mente en términos de mecanismos computacionales y representaciones mentales. Esta parte de la tesis es la que recoge correctamente la TRCM. Sin embargo, del hecho de que se postule que la mente posee un sistema representacional no se sigue, como los filósofos de la psicología creen, que "la idea de que la mente es un sistema representacional es por tanto la idea de que la mente es un sistema intencional" (Flanagan 1984, p. 178). La asimilación tiene lugar porque dado que el contenido de los estados intencionales es acerca del mundo y dado que es en virtud de tener contenido que los estados mentales son representacionales, la noción de contenido hace las veces de puente para la asimilación entre intencionalidad y representación.

El problema básico para estos filósofos es explicar el fenómeno de la intencionalidad y para ello se postula un sistema representacional interno. Se puede acordar con la idea de que los estados intencionales son estados representacionales pero eso no implica aceptar que todo estado representacional es intencional, en este sentido el ámbito de la representación es mucho más amplio que el de la intencionalidad (dado que habría estados que representan pero que no serían acerca

de nada en el mundo). Incluso, se puede sostener que el fenómeno de la intencionalidad tiene una relación tan remota (en el sentido de sumamente compleja) con el fenómeno de la representación que hasta sería conveniente tratarlos en ámbitos diferentes. De todo esto me ocuparé en los capítulos 4 y 5.

Con respecto a la tesis 1, el objetivo de la psicología cognitiva es explicar los mecanismos internos (los estados y procesos mentales) que subyacen a (o causan) la conducta.<sup>35</sup> En cambio, la idea de los filósofos de la mente es que el objetivo de las explicaciones psicológicas es explicar la conducta intencional de los individuos y es para ello que se postulan estados intencionales.<sup>36</sup> Así, la psicología tiene como objetivo la explicación de la acción intencional en términos de deseos y creencias (o de APs). Esto aparece claramente en Fodor (1968) y en los escritos de innumerables filósofos de la psicología.

Pocos filósofos rescatan la tesis 1 afirmando que “se podría pensar en el psicólogo cognitivo como interesado en ofrecer una descripción funcional abstracta de la mente” (Flanagan 1984, p. 180), o que el “objetivo es la construcción de una teoría científica acerca de los mecanismos que subyacen a la conducta” (Stich 1983, p. 6). En estas definiciones no se alude a que el objeto de estudio sea la acción intencional ni que los estados pertinentes sean los de AP, tal como el realismo intencional los entiende sino que se alude a mecanismos internos descritos funcionalmente. En este sentido, algunos autores afirman explícitamente que “Fodor, debido a que se involucró tempranamente en la investigación del procesamiento de oraciones es tomado seriamente por muchos psicolingüistas... sin embargo... su prescripción de que la psicología cognitiva es principalmente acerca de las actitudes proposicionales es ampliamente ignorada” (Garnham 1994, p. 175). La cuestión es, como dije antes, que no sólo Fodor cree que la tarea de la psicología es

---

<sup>35</sup> O en términos más ambiciosos: “la psicología cognitiva intenta comprender la naturaleza de la inteligencia humana y cómo piensa la gente” (Anderson 1980, p. 3).

<sup>36</sup> Tal como mencioné en la nota 2 de la introducción, voy a hablar indistintamente de conducta y acción para referirme a la descripción intencional. Esto se debe a que la distinción que realmente interesa es entre la conducta o acción intencional, y los procesos internos (o la distinción entre c y d de la nota 2 mencionada). Cuando sea necesario hacer la distinción entre conducta (movimientos corporales públicamente observables) y acción intencional (descripción de la conducta en términos intencionales que aluden a las creencias y deseos de la persona cuya acción se quiere explicar), la haré explícita.



dar cuenta de la acción en términos de APs sino todos los filósofos que de alguna u otra forma adhieren a la TRCM (e incluso muchos que no adhieren a la misma), no olvidar que Fodor la postuló como aquella teoría que subyace a todas las teorías psicológicas. Desarrollaré este punto, en particular, en los capítulos 4 y 5.

Por último, con respecto a la tesis 3, y muy relacionado con lo que se viene diciendo, la psicología suele adoptar las explicaciones funcionales y no las explicaciones intencionales. Es habitual la presentación de las teorías psicológicas en términos de diagramas de cajas en las que se muestra cómo una capacidad puede descomponerse en subcapacidades más simples que explican la capacidad analizada. El análisis funcional no está pensado para capacidades intencionales sino semánticas (aquellas en cuya descripción no intervienen las APs). Esto quedará claro, en especial, en el capítulo 5.

En conclusión, la TRCM soluciona el problema de la intencionalidad incorporando la PSC a la psicología científica a costa de interpretar erróneamente dos y media de las tres tesis básicas de la psicología cognitiva. Sin embargo, la TRCM por sí sola no resuelve un aspecto fundamental del problema de la intencionalidad que es el de explicar cómo los estados físicos poseen la propiedad de ser intencionales. Esto lo veremos en el capítulo 3 cuando analice el proyecto naturalista de una semántica intencional para las representaciones mentales postuladas por la TRCM (y por ende, para las postuladas por las teorías psicológicas dado que se supone que la TRCM es la teoría subyacente a las mismas). Pero antes, en el capítulo 2, seguiremos viendo cómo la filosofía de la psicología asimila el problema de la representación al de la intencionalidad examinando algunas cuestiones en torno a las representaciones y a la TRCM.

## 2

## El problema de las representaciones y la teoría representacional de la mente

En el capítulo anterior hemos visto cómo las cuestiones acerca de las representaciones mentales y las cuestiones acerca de la intencionalidad se entremezclan desde el nacimiento mismo de la filosofía de la psicología cognitiva en la TRCM. En este capítulo seguiré desarrollando este punto con mayor detalle. En la primera parte, presento ciertas cuestiones básicas en torno a las representaciones que toda teoría que se ocupe del tema debería abordar (§1.1 a §1.5). A lo largo del desarrollo de estas cuestiones comenzaré a desligar el ámbito intencional del representacional. En la segunda parte, reconstruyo las tesis básicas de la teoría mayoritariamente aceptada por los filósofos de la psicología como aquella que subyace a las teorías psicológicas: la Teoría Representacional de la Mente (TRM) (§2.1) y presento las respuestas de esta teoría a las cuestiones básicas desarrolladas en la primera parte (§2.2).

### 1. El problema de las representaciones

Cualquier acercamiento al problema de las representaciones mentales debe tener en cuenta, al menos, las siguientes cuestiones:

1. Los criterios para postular representaciones
2. Los requisitos para una teoría de la representación

3. El formato de las representaciones
4. La implementación de las representaciones
5. La definición de representación

La primera cuestión se relaciona con la idea de que no cualquier organismo o artefacto requiere poseer, o es necesario atribuirle, representaciones mentales que medien entre sus inputs y sus outputs. De la misma manera, no todas las funciones de los subsistemas que conforman un organismo o un artefacto hacen uso de representaciones para llevar a cabo su tarea. Postular representaciones requiere alguna noción de sistema cognitivo y para que algo sea (o se le atribuya ser) un sistema cognitivo se requieren ciertos requisitos. La segunda cuestión también alude a ciertos requisitos pero en este caso son requisitos que debe cumplir cualquier teoría acerca de las representaciones mentales de los sistemas cognitivos. La tercera se refiere al formato que adoptan las representaciones, esto es, si son imágenes, símbolos, estados neurofisiológicos y/o vectores de activación. La cuarta alude a cómo se implementan (realizan, instancian) las representaciones mentales en un sistema físico. La quinta cuestión tiene que dar una respuesta a qué es para algo representar y qué es para algo representar algo específico.

De las cinco cuestiones, algunas son claramente empíricas como 3 y 4, y otras claramente conceptuales como 2 y 5. La 1 requiere de ciertos desarrollos conceptuales a la vez que de cierta investigación empírica. De todas formas, la distinción entre la investigación empírica y la tarea conceptual no tiene un límite preciso y no es mi intención adherir a ningún criterio demarcatorio. Así, la cuestión de la implementación necesita investigación empírica pero al mismo tiempo hacen falta teorías acerca de la implementación. Lo mismo ocurre con el formato representacional. Lo que quiero señalar es que no se puede dilucidar estas cuestiones haciendo sólo filosofía de la psicología. Por otro lado, una teoría acerca de qué es representar está constreñida por cuestiones empíricas acerca del formato y la implementación de las representaciones. En última instancia, las distintas cuestiones actúan como restricciones entre sí y 1 a 4 restringen fuertemente lo que sea representar.

En lo que sigue, se verá en qué consiste cada problema, analizaré algunas de las respuestas propuestas e iré restringiendo el ámbito de la discusión que desarrollaré en los próximos capítulos. Dado que las cuestiones 1 a 5 han sido elaboradas en el marco de la filosofía de la psicología que asimila estados representacionales a estados intencionales, gran parte de las mismas y las respuestas que han recibido, han tenido como objetivo el problema de la intencionalidad. Desarrollaré 1 a 5 tal como han sido planteadas y marcaré la distinción en donde sea pertinente.

### *1.1. Los criterios*

El primer problema consiste en que no cualquier organismo o artefacto (o sus subsistemas) requiere poseer, o es necesario atribuirle, representaciones mentales para dar cuenta de sus funciones. Si tomamos la noción de sistema cognitivo, parecería haber una distinción intuitiva entre sistemas como las personas, y los organismos como las paramecias o los artefactos como los termostatos. Por otro lado, en el caso de los organismos humanos, parece intuitivo y obvio pensar que para dar cuenta de la función del sistema digestivo no hace falta atribuirle a ese subsistema la manipulación de representaciones, en cambio parecería adecuado atribuir representaciones si de lo que se trata es de dar cuenta de ciertas capacidades como el pensamiento. Dado que la postulación de representaciones fue ideada para sistemas (o subsistemas) cognitivos, se requieren ciertas condiciones que un sistema satisfaga para ser considerado cognitivo.

La razón primordial (y hasta me animaría a decir que es la única) para postular representaciones radica en una inferencia a la mejor explicación (cf. Fodor 1975, Pylyshyn 1984). Dados ciertos fenómenos a explicar y dado que de todas las explicaciones en competencia la explicación que postula representaciones mentales es la mejor, según ciertos criterios de elección entre explicaciones, es probable que esta explicación sea la adecuada. Los criterios de elección estarían dados básicamente por el poder explicativo de la teoría que postula representaciones

mentales, así la idea sería que no parece haber ninguna explicación que no las postule que pueda dar cuenta de ciertos fenómenos. De manera que dado este criterio metodológico para atribuir representaciones a un sistema cuyas capacidades se intentan explicar, si se puede obtener una mayor comprensión de cómo funciona describiendo partes del mismo como representaciones, ésta es la mejor evidencia que se puede obtener de que el sistema tiene representaciones (van Gelder 1995). Así, desde un argumento metodológico se llega a uno ontológico acerca de la existencia de representaciones mentales.

El argumento que funda la inferencia a la mejor explicación se basa en varios argumentos que se podrían agrupar en dos tipos. Uno relacionado con la distinción entre sistemas cognitivos y cosas tales como rocas, el cielo, los termostatos y las paramecias, que no son sistemas cognitivos. Y el otro, relacionado con la diferencia al interior de los sistemas cognitivos entre capacidades cognitivas y otras que no lo son.<sup>1</sup> Veamos primero el segundo tipo. El argumento es el conocido acerca de la pobreza del estímulo. Este argumento ha sido utilizado para postular tanto representaciones (Chomsky 1965, Fodor 1975) como mecanismos mentales (Fodor 1983). En el caso de las representaciones, dada la pobreza de los estímulos, la única manera de explicar la riqueza y diversidad de los outputs conductuales es postulando un sistema representacional interno.<sup>2</sup>

Este argumento se basa en la distinción entre dos tipos de explicaciones de la conducta: la mecánica y la intencional (Dennett 1987, Bermúdez 1995). Las explicaciones mecánicas postulan relaciones nomológicas que conectan estímulos con respuestas. Las respuestas son invariables producto del condicionamiento asociativo (como en los experimentos con ratas), o porque dichas conexiones están cableadas en el sistema. Dado un determinado input, el rango del output es sumamente limitado (o casi único) y la especificación del output se hace en términos de movimientos

---

<sup>1</sup> Por supuesto que un argumento a favor de lo primero puede constituir un argumento a favor de lo segundo: si hay criterios de distinción entre sistemas cognitivos y otros que no lo son, seguramente por las mismas razones es posible establecer, al interior de los sistemas cognitivos, la diferencia entre subsistemas cognitivos y otros que no lo son. Esto es lo que a algunos les permite sustentar la idea de que hay mecanismos y estados que compartimos con sistemas no cognitivos como los animales no superiores (e incluso con algunos artefactos).

<sup>2</sup> En paralelo, la única manera de dar cuenta de la diferencia en riqueza del input y el output es postulando mecanismos especializados según el tipo de input.

corporales no-relacionales, esto es, descripciones que no apelan a las causas distales de los estímulos. En cambio, las explicaciones intencionales no involucran conexiones nomológicas estímulo-respuesta ya que es posible su disociación. Hay respuestas que ocurren en ausencia de un estímulo (por ejemplo, producto de la imaginación) y hay ocasiones en las que se registran estímulos sin que haya una respuesta (por ejemplo, cuando captamos ciertos ruidos a los que no respondemos). Dado un determinado input, el rango del output es sumamente amplio y se especifica en términos relacionales, esto es aludiendo a las causas distales de los estímulos.

Dada la distinción entre estos tipos de explicación la idea es que hay ciertas capacidades y conductas para las cuales no hay una simple explicación mecánica y por ende, es teóricamente inevitable apelar a estados representacionales, esto es, estados intermedios entre los inputs y los outputs. Otra manera similar de establecer la diferencia es apelando a la distinción personal-subpersonal (Dennett 1969, pp. 90-6; 1987, cap. 3). Hay una diferencia entre las explicaciones neurofisiológicas abstractas de los subsistemas cognitivos que procesan información interna al, y entre partes del, organismo (la máquina sintáctica) y la explicación que da cuenta de la conducta del organismo, como un todo, en términos intencionales (la máquina semántica).

El cerebro en tanto máquina sintáctica procesa información que ingresa a través de los transductores (perceptivos) que traducen el input al formato sobre el que operan los mecanismos neurofisiológicos y luego traducen (los transductores motores) los resultados de los mecanismos en conductas. Estos mecanismos operan sobre las propiedades estructurales (sintácticas) de los estímulos de manera que no tienen acceso al contenido de la información. Tampoco la persona (el sistema global) tiene acceso al contenido de la información que procesan los mecanismos subpersonales, de modo que sólo en el nivel personal los estados internos son significativos para el sistema (global). Por eso, sólo tiene sentido atribuir estados intencionales a la persona, son los estados personales los que tienen contenido intencional. El cerebro en tanto máquina semántica discrimina significados y la persona actúa conforme a ellos, esto es, conforme a sus estados internos intencionales. La relación entre ambos niveles consiste en que los mecanismos

subpersonales explicarían cómo el cerebro implementa las descripciones del nivel intencional.

En mi opinión, la primera distinción entre explicaciones mecánicas e intencionales (aunque sería conveniente denominar a estas últimas 'explicaciones que apelan a representaciones mentales') recoge adecuadamente la distinción entre ciertos mecanismos para cuya descripción no es necesario apelar a estados representacionales como por ejemplo la digestión, de aquellos para los cuales es inevitable apelar a estados representacionales como las capacidades perceptivas y lingüísticas, y las superiores, como el pensamiento y la resolución de problemas. En cambio, tal como entenderé la distinción personal – subpersonal (ver el capítulo 5 y la conclusión), ésta es una distinción al interior del nivel representacional. El nivel subpersonal es un nivel tan cognitivo como el personal y en ese sentido las descripciones de ambos niveles apelan a estados representacionales. Por ejemplo, tanto para la capacidad lingüística como para la visión, Chomsky y Marr han propuesto explicaciones que están en el nivel subpersonal. Sin embargo, estas explicaciones no son en términos de movimientos corporales o neurofisiológicos sino en términos psicológicos, esto es, que aluden al ámbito de lo cognitivo.

Es por esto que no acuerdo con la definición de Dennett en la que lo subpersonal está exclusivamente ligado con movimientos corporales y eventos en el sistema nervioso; y con aquellos que asimilan la distinción mecánico-representacional con la subpersonal-personal (por ejemplo, Bermúdez 1995) haciendo equivaler lo subpersonal con lo mecánico y lo personal con lo representacional o cognitivo. El nivel subpersonal es un nivel cognitivo y en ese sentido hay estructura mental, no es meramente mecánico. En el capítulo 5, donde exploro las consecuencias para la psicología cognitiva que se siguen de mi propuesta de un naturalismo amplio (capítulo 4), me ocupo en particular de la capacidad lingüística subpersonal y allí quedará claro por qué el nivel subpersonal es cognitivo.

El problema con ambas distinciones es que proponen un criterio de demarcación en el que uno de los ámbitos es el nivel intencional que colapsa con el ámbito representacional. Ya hemos visto algo en torno a esta confusión y ésta quedará más clara cuando veamos en la segunda parte de este capítulo la teoría

representacional de la mente. En el capítulo 5 volveré sobre la distinción personal – subpersonal de manera que espero que quede claro que la distinción intenta diferenciar dos tipos de contenido representacional, aquel puramente interno al sistema (el ‘contenido interno’ o contenido representacional propiamente dicho) y aquel para cuya individuación se requiere apelar no sólo a otros estados internos del individuo sino al entorno físico y social (el ‘contenido intencional’).

Pasando a la cuestión de la distinción entre sistemas cognitivos y aquellos que no lo son, habría, en principio, dos estrategias, una realista y otra instrumentalista. La pregunta para la filosofía de la psicología en vez de ser ¿qué es lo que distingue a un sistema cognitivo de uno que no lo es? es ¿qué es lo que hace que un sistema sea intencional? De todas formas, intentaré extraer de la respuesta a la segunda pregunta algunos criterios que sirvan para responder la primera pregunta. Veamos primero la estrategia instrumentalista.

Dennett (1987) propone tratar a aquellos sistemas de los cuales queremos predecir su comportamiento como sistemas intencionales. Un sistema es intencional si le es apropiada una actitud intencional. Una actitud intencional le es apropiada a un sistema si podemos hacer predicciones adecuadas a partir de ciertos patrones conductuales. La predicción del comportamiento se realiza en virtud de la atribución de racionalidad, creencias y deseos al sistema en cuestión. Es decir, se atribuyen los estados intencionales que un sistema debería tener, dadas las circunstancias y los objetivos del mismo, y se predice que se comportará de manera racional, esto es, en virtud de tales estados intencionales.

La actitud intencional es una de las estrategias que adoptamos los seres humanos, además de la actitud de diseño (la predicción en base al fin para el que un sistema está diseñado o apelando a una descripción de su funcionamiento) y la actitud física (la predicción en base a la determinación de la constitución física del sistema utilizando las leyes físicas). Sin embargo, la actitud intencional es, al menos para los seres orgánicos, más básica que la de diseño porque esta última supone adoptar la primera (supone entender los procesos evolutivos de diseño en términos intencionales, cf. Dennett 1993, p. 224).



El criterio pragmático que avala la adopción de la actitud intencional (puesto que no habría ninguna restricción 'de principio' para considerar a cualquier sistema como intencional) consiste en el poder predictivo de la misma. La diferencia entre los sistemas genuinamente intencionales y los que no lo son, no consiste en una diferencia en la constitución interna o en la complejidad del diseño sino en el fenómeno objetivo del éxito o fracaso de la actitud intencional. Hay modelos que describen el comportamiento humano desde la actitud intencional y, que permiten obtener generalizaciones y predicciones de gran poder informativo para la vida cotidiana. De manera que, parece ser que sólo aquellos sistemas para los que hay modelos del comportamiento basados en la actitud intencional (seres humanos, ciertos animales y computadoras) pueden ser considerados sistemas intencionales.

Fodor (1986) propone un criterio realista que permite afirmar que la teoría representacional de la mente es verdadera para aquellos organismos a los cuales se aplica. El criterio es la posesión por parte de esos organismos de una propiedad que consiste en responder selectivamente a las propiedades no-nómicas de los estímulos. Una propiedad nómica es aquella que participa de leyes, de manera que los objetos que la poseen están subsumidos bajo esas leyes en virtud de esa propiedad. Todo organismo puede responder selectivamente a propiedades nómicas que son detectadas por transductores, sin embargo sólo algunos sistemas pueden responder selectivamente a propiedades no-nómicas que no pueden ser detectadas por transductores.

Ciertas propiedades como ser una camisa arrugada, no forman parte de leyes y por ende ningún objeto cae bajo leyes en virtud de ser una camisa arrugada, pero sin embargo la conducta humana puede responder selectivamente a esta propiedad. La idea es que un objeto tiene varias propiedades nómicas y no-nómicas, estas últimas causan (por medio de relaciones nómicas psicofísicas) ciertos estados mentales en un organismo. Este infiere de las propiedades nómicas detectadas por los transductores la propiedad no-nómica de manera que el organismo representa *S como O* (la camisa en tanto arrugada). La información de que S tiene ciertas propiedades nómicas sirve como premisas para una inferencia perceptiva de que S tiene la propiedad no-nómica O (por ejemplo, de ciertos patrones lumínicos el individuo infiere que la camisa tiene

la propiedad de estar arrugada). Dicha representación está nómicamente conectada con la conducta específica del organismo. Así, lo que conecta que S sea O con la respuesta específica del organismo es una representación mental que representa S como O.<sup>3</sup>

Ya sea que se adopte un criterio realista o pragmático parece que las diferencias en capacidades conductuales entre las personas y otros organismos o artefactos están dadas por la capacidad racional o inferencial.<sup>4</sup> Así, una condición necesaria para que algo sea intencional es que su comportamiento sea racional, esto es, sea acorde a las representaciones mentales que causan o son razones para la conducta. Ni el realista ni el instrumentalista con respecto a la postulación de representaciones mentales dudan que los seres humanos son racionales bajo alguna concepción apropiada de racionalidad. Sin embargo, no hace falta comprometerse con ninguna concepción en particular sino simplemente con la idea de que el dominio de lo cognitivo está marcado por ciertos principios de inferencia que permiten ir desde un determinado input a un determinado output.

Los sistemas cognitivos instancian funciones cognitivas cuyos argumentos y valores están relacionados epistémicamente, esto es, dado ciertos inputs, el output es racional (en el sentido de 'convinciente') a la luz de aquellos (Cummins 1983, cap. 3;

---

<sup>3</sup> Hay otros intentos de dar criterios realistas para la distinción. En particular, aquel que alude a la diferencia entre la intencionalidad intrínseca (o genuina) y la derivada (Searle 1983). Se dice que la intencionalidad genuina es la de los estados mentales de los humanos y lo es en dos sentidos, en tanto que es una propiedad intrínseca (no-relacional y en ese sentido superviene a los estados físicos del individuo) y en tanto que no depende de la interpretación de un observador. La intencionalidad de las computadoras o demás artefactos y animales es derivada de la de los humanos en el sentido de que estos últimos le atribuyen intencionalidad (por ejemplo, al decir que la computadora no 'quiere' funcionar) pero la atribución es metafórica ('como-si'). Descarto la aplicación de este criterio para los sistemas cognitivos no sólo porque es completamente *a priori* sino porque cae completamente afuera de una concepción funcionalista de la mente.

Otros criterios que apuntan al dominio de los problemas para los cuales se postulan representaciones, a la complejidad de los parámetros del entorno y a requisitos del sistema representacional interno se pueden encontrar en Haugeland (1991), p. 62, y Clark y Toribio (1994). En ambos, aunque más explícitamente en el último, se requieren capacidades inferenciales.

<sup>4</sup> Cabe hacer notar que mientras que en Fodor (1986) las propiedades intencionales son propiedades no-nómicas, ya en Fodor (1987) y más claramente en Fodor (1994), estas propiedades pasan a ser nómicas, esto es, se postula la existencia de leyes intencionales. Con lo cual la distinción de (1986) entre sistemas cognitivos y no cognitivos se diluye. Sin embargo, es la cuestión de la descripción de las capacidades cognitivas como inferenciales lo que interesa para marcar la distinción entre sistemas cognitivos y sistemas que no lo son, aunque esta distinción no se base en la diferencia entre propiedades nómicas y no-nómicas.

1989, cap. 8; Egan 1992, p. 452).<sup>5</sup> Así, los sistemas cognitivos son mecanismos inferenciales que toman argumentos y valores y los relacionan como lo están las premisas con la conclusión. Los inputs y los outputs tienen una interpretación proposicional dado que las proposiciones son el tipo de elementos que participan de procesos inferenciales y en ese sentido, están descriptos en términos semánticos.<sup>6</sup> Todo sistema cognitivo procesa información y este procesamiento puede verse como transformaciones de instancias de oraciones, en este sentido toda capacidad cognitiva puede caracterizarse como una capacidad inferencial. En general, tanto para mecanismos subpersonales, por ejemplo, el procesamiento visual temprano o el análisis lingüístico, como para los personales, por ejemplo, el pensamiento, hay una descripción en términos inferenciales.<sup>7</sup>

Ofrecer una descripción de un sistema cognitivo como aquél en el que ciertas capacidades son vistas como inferenciales es ofrecer (parte de) una propuesta de caracterización de la organización funcional de un sistema. Hemos visto en el capítulo anterior que tener una mente es tener una cierta organización funcional. Establecer cuál es la organización funcional completa y apropiada de los sistemas cognitivos es una tarea empírica pero, en principio, se puede pensar que todo sistema cognitivo tiene que tener (al menos) una organización funcional lo suficientemente similar a la que se acaba de describir. El chauvinismo y el liberalismo son los problemas básicos a los que se enfrenta cualquier caracterización funcional. Describir (en parte) a los sistemas cognitivos como sistemas inferenciales no parece ser tan restrictivo como para que sólo pueda aplicarse a nosotros (chauvinismo) ni tan laxo como para que incluya sistemas que no quisiéramos incluir como cognitivos (liberalismo). Adoptaré esta concepción habitual de la organización funcional de un sistema cognitivo como

---

<sup>5</sup> Uso "convinciente" (o también podría ser "persuasivo") para el término inglés "*cogent*". La idea es que una secuencia es convincente cuando cada parte lleva convincentemente a la siguiente, de manera que cada consecuencia es la continuación natural e inevitable de su antecedente.

<sup>6</sup> Que haya una interpretación de los inputs y los outputs en términos proposicionales no implica ningún compromiso en cuanto al formato representacional, esto es, no implica un compromiso ontológico con símbolos.

<sup>7</sup> Tanto Marr (1982) como Chomsky (1980b) describen el procesamiento de las respectivas capacidades subpersonales como inferencial. Si hubiera capacidades no describibles de manera inferencial, de todas formas lo serían al estar en sistemas cognitivos inferenciales (Cf. Cummins 1983, cap. 3). Por otro lado, no hay ningún filósofo de la mente que no conciba las capacidades

aquel cuyas capacidades pueden interpretarse como capacidades inferenciales y cuyos estados pueden describirse en términos semánticos.

### *1.2. Los requisitos*

La segunda cuestión se refiere a los requisitos que debe cumplir cualquier teoría de las representaciones mentales. Si se hace un repaso por la literatura sobre el tema, los requisitos que se han propuesto son, básicamente, para una teoría del contenido intencional.<sup>8</sup> Veamos en qué consisten de manera de establecer cuáles serían pertinentes para una teoría de las representaciones mentales:

(a) *requisito explicativo*: la teoría debe sustentar, o al menos, ser compatible con el rol explicativo que las ciencias cognitivas le otorgan a las representaciones.

(b) *requisito implementacional*: la teoría debe ser compatible con las mejores explicaciones científicas acerca de qué entidades cumplen el rol representacional en la mente/cerebro.

(c) *requisito naturalista*: la teoría no debe apelar a términos semánticos o intencionales para dar cuenta de la noción de representación.

(d) *requisito del contenido intencional*: la teoría debe explicar las atribuciones de contenido intencional.

(e) *requisito de la representación errónea*: la teoría debe dar cuenta de cómo es posible que a veces las representaciones representen lo que no es el caso.

---

personales que hacen uso de APs como inferenciales. Esto se ve claramente en las explicaciones intencionales.

<sup>8</sup> Véase, en particular, Warfield y Stich (1994), y Bermúdez (1995).

(f) *requisito de la integración cognitiva*: la teoría tiene que explicar cómo las representaciones se integran con otras representaciones.

No es el objetivo de esta tesis ofrecer una teoría de las representaciones mentales sino la crítica a un aspecto de la teoría representacional de la mente, aquel que consiste en el proyecto de una semántica para las representaciones. Como consecuencia de esta crítica quedará esbozada en los últimos capítulos una manera de entender las representaciones que participan de mecanismos y estados psicológicos. Este esbozo de propuesta responde a la distinción entre el ámbito de la representación y el ámbito de la intencionalidad, de modo que un adecuado desarrollo de la misma cumpliría con algunos de los requisitos enumerados más arriba, aquellos que no aluden al ámbito intencional.

En particular el requisito (d) y (e) están arraigados a la idea de que una teoría de las representaciones mentales es una teoría de la intencionalidad. El requisito (d) afirma que la noción de representación debe ser lo suficientemente de grano fino como para recoger los rasgos típicos de los estados intencionales y para participar en las explicaciones típicas de la atribución intencional. El rasgo típico de los estados intencionales es la modularidad proposicional que consiste en tres propiedades (Ramsey, Stich y Garon 1990). En primer lugar, el ser funcionalmente discretos, es decir, se puede hablar de agentes que obtienen o pierden creencias individuales. En segundo lugar, el ser interpretables semánticamente, esto es, tienen propiedades semánticas tales como condiciones de satisfacción. Las propiedades semánticas permiten hacer generalizaciones nomológicas que involucran contenidos que figuran en los *explananda* de las explicaciones de la conducta de los individuos que tienen la forma típica de 'Si un individuo desea *X*, y cree que no puede lograr *X* a menos que haga *Y* entonces, *ceteris paribus*, intentará hacer *Y*. Y, finalmente, son causalmente eficaces, esto es, producen otros estados intencionales y conducta.

Cómo se verá a lo largo de esta tesis, no considero que cuestiones acerca de la intencionalidad o la atribución intencional participen de la explicación y comprensión de las capacidades psicológicas propiamente dichas, aquellas del nivel

subpersonal (al menos, de las capacidades ampliamente estudiadas como el lenguaje y la visión). Las cuestiones acerca de la intencionalidad se relacionan con el problema de cómo nuestras mentes representan de manera correcta cómo es el mundo de modo de comportarnos acorde con esas representaciones. Las representaciones de las que se habla aquí son las que participan de los estados intencionales y el problema es cómo nuestros estados intencionales están conectados con el mundo.

No creo que sea la tarea de la psicología cognitiva dar cuenta de cómo nuestras creencias, deseos, y demás se relacionan con el entorno y en ese sentido, una teoría de las representaciones no tiene por qué ceñirse al requisito (d). Explicar cómo un organismo representa verdaderamente su entorno, o el éxito de las interacciones del organismo con su entorno depende de cuestiones que van más allá de una investigación psicológica que se circunscribe a mecanismos subpersonales. De manera que la corrección o incorrección de las representaciones no es un asunto que sea pertinente a la hora de dar cuenta del funcionamiento de ciertos mecanismos, los mecanismos funcionan igual ya sea que las representaciones sean falsas o verdaderas. Por eso, no considero que el requisito (e) sea relevante para una teoría acerca de las representaciones que intervienen en los mecanismos psicológicos.

Con respecto al resto de los requisitos, el (a) es el más importante. Una teoría de las representaciones debe ser compatible con el rol explicativo que los psicólogos cognitivos le otorgan a las representaciones. Este es el requisito que guía la propuesta del capítulo 5 basada en el rol explicativo que ciertas teorías le otorgan a las representaciones (en particular, la lingüística cuyo precursor fue Chomsky y la teoría de la visión cuyo precursor fue Marr). También (b) es importante a la hora de ponderar cuáles son los vehículos de las representaciones, en ese sentido, mi esbozo de propuesta no se compromete con ninguna hipótesis en particular, dado que será la investigación empírica quien establecerá qué cosas cumplen el rol representacional en la mente/cerebro.

Por otro lado, toda teoría de las representaciones debe poder dar cuenta de (f). Un aspecto de la integración cognitiva radica en cómo las representaciones que

intervienen en los distintos subsistemas que componen una determinada capacidad se integran de manera de dar lugar a la misma. Otro aspecto, y el más complicado, es cómo las representaciones que participan en capacidades como el pensamiento se integran con las de los estados perceptivos, motivacionales, de atención, y demás, y con representaciones almacenadas, para la coordinación de distintas capacidades. Ambos aspectos podrían verse como la cuestión de las relaciones entre el nivel subpersonal y el personal. Visto desde esta perspectiva, esta cuestión es sumamente compleja y depende de cómo se entienda la distinción. Hasta donde sé, nadie ha ofrecido un enfoque que dé cuenta de las relaciones entre los mecanismos y estados de ambos niveles, aunque se han delineado propuestas en favor de que tiene que haber algún tipo de conexión.<sup>9</sup>

El requisito naturalista (c) tiene motivaciones ontológicas y metodológicas. Las ontológicas radican en que dado que las propiedades semánticas o intencionales no son propiedades básicas del universo, deben poder explicarse en función de propiedades básicas como las físicas o las que las ciencias naturales establezcan que hay. Las metodológicas radican en evitar una definición circular de la noción de representación en términos semántico/intencionales dado que la noción misma de representación mental es una noción semántica/intencional. La idea que sustenta este requisito es creer que si no hay un enfoque naturalista de los estados intencionales, estos no formarían parte del mundo natural y por ende, quedarían eliminados *qua* estados intencionales.

Dado que este es uno de los temas centrales que se abordará en los capítulos 4 y 5 no diré mucho al respecto. Simplemente, que no es tanto el requisito de naturalización de la noción de representación el que debería cumplir toda teoría sino el de abordar su objeto de estudio de una manera naturalista. Este naturalismo implica abordar el estudio de las personas en tanto inmersas en el mundo natural. No son en particular las representaciones las que deben conectarse con el mundo para ofrecer una explicación naturalista de los aspectos mentales de las personas sino que las personas en tanto tales ya están inmersas en un mundo natural. Un

---

<sup>9</sup> Para esto último véase, por ejemplo, Bermúdez (2000) y Davies (2000). Aunque, véase también la nota 48 del cap. 5.

mundo natural que abarca tanto el entorno físico como social. En este sentido, es posible dar cuenta de la intencionalidad de manera naturalista pero sin comprometerse con un proyecto de naturalización de las representaciones mentales. Este es uno de los aspectos centrales de esta tesis.

### *1.3. El formato*

La tercera cuestión se relaciona con el formato o vehículo de las representaciones, si son, por ejemplo, imágenes, símbolos, estados neurofisiológicos, o vectores de activación. Los que consideran que son los estados neurofisiológicos los que tienen un rol representacional, en general creen que o bien la cuestión de las representaciones mentales es un fenómeno puramente biológico y en ese sentido no hay un nivel intermedio de descripción entre el del sentido común y la neurofisiología (Searle) o bien que no hay manera de reducir el nivel mental al físico y por ende, en una psicología científica el nivel de descripción mental queda eliminado (los Churchland).

La idea de que las representaciones son imágenes que se asemejan o copian lo representado ha sido sostenida históricamente por los empiristas ingleses Locke y Hume. Debido a que la relación de semejanza ha recibido innumerables críticas, esta teoría cayó en desuso. Actualmente, los defensores de esta postura sostienen que las representaciones representan en el mismo sentido en que las imágenes representan. Según Kosslyn (1980, 1994), una representación mental es una imagen funcional del objeto representado, la representación no es una imagen real sino que es 'como si' fuera una imagen, como si poseyera características espaciales que se corresponden con las de los objetos representados.

Pylyshyn (1981a, 1981b) sostiene que las imágenes mentales representan en el sentido en que lo hacen las descripciones lingüísticas (en un código neural). Las imágenes son descripciones estructurales explícitas de propiedades de los objetos representados y en este sentido, el formato representacional de las imágenes sería simbólico con propiedades semánticas y sintácticas a la manera de un lenguaje. Por



otro lado, el simbolismo del cognitivismo clásico se contrapone a una corriente del conexionismo que sostiene que hay que entender a las representaciones como vectores de activación que propagan algoritmos sobre una red de nodos. Si bien estos vectores tienen un rol representacional y en ese sentido pueden describirse en términos semánticos, no tienen la estructura constitutiva de un lenguaje a la manera de los constituyentes sintácticos del cognitivismo clásico (Smolensky 1988, 1991).

Dado que no me comprometo con ningún formato en particular, no me ocuparé de esta cuestión. Pudiera ser el caso que tuviéramos distintos formatos representacionales dependiendo del nivel de la capacidad cognitiva (si es del nivel personal o subpersonal) o del tipo de capacidad cognitiva que se ponga en juego en cualquiera de esos niveles. En algunos casos, por ejemplo para algunas funciones cognitivas superiores (o personales o intencionales, como el pensamiento), daré por sentado (parte de) lo que las teorías representacionales de la mente sostienen y es que las representaciones son simbólicas.

Sin embargo, en relación con la noción de representación que interviene en las capacidades subpersonales (cuestión de la que me ocupé en el capítulo 5) no me comprometo con ningún enfoque acerca del formato. Esto es así, básicamente porque sólo es un esbozo de noción. No obstante, cualquier teoría de las representaciones mentales desarrollada en detalle tendrá que comprometerse de alguna manera u otra con alguna de las ofertas acerca del formato dado que, tal como veremos en la segunda parte de este capítulo, parece que el formato carga en sí algunas de las propiedades que permitirían dar cuenta de ciertos hechos acerca de nuestras mentes (por ejemplo, la productividad y la sistematicidad del pensamiento).

#### *1.4. La implementación*

La cuarta cuestión es la de la implementación, esto es, cómo se implementan o realizan o instancian las representaciones mentales en el cerebro. Para dar cuenta

de esto hay que ofrecer una teoría acerca de la implementación. Hemos visto que una de las tesis básicas de la ciencia cognitiva es que existe un nivel de descripción de la función de un sistema cognitivo que es especificable en términos de procesamiento de la información y que este nivel se entiende en términos computacionales. Así, que un sistema computacional sea una máquina de sumar radica en que puede ser descrito por una 'función +':  $+( \langle n, m \rangle ) = s$ .<sup>10</sup> Esta 'función +' toma como argumentos (o inputs) la secuencia de botones apretados:  $\langle C, N, +, M, = \rangle$ , esto es,  $\langle$ despejar, primer adendo, más, segundo adendo, igual $\rangle$  y como valores (o outputs) los numerales exhibidos en una pantalla P:  $n + m$ . La suma relaciona números, eso significa que el sistema físico no satisface literalmente la 'función +' sino que lo que hace es instanciar la 'función +' satisfaciendo la 'función g' (los argumentos y valores de g son estados físicos del sistema). Bajo interpretación I, la función g es una instanciación de la función +.

Esta 'función de interpretación' relaciona el ámbito simbólico computacional con un dominio representado. En cambio, la relación entre el ámbito computacional y el físico está dada por una 'función de realización' que proyecta estados simbólicos en estados físicos del sistema (Egan 1992). Además de las relaciones entre los distintos niveles de una teoría psicológica (nivel computacional, algorítmico y de implementación, según Marr 1982) en términos de la función que es instanciada, una teoría de la implementación requiere dar cuenta de la coinstanciación de las propiedades cognitivas y computacionales (o físicas). Es decir, necesita dar cuenta no sólo de cómo se instancia una función sino de por qué determinados estados representacionales se instancian en determinados estados computacionales o físicos.

La ciencia cognitiva hace hincapié en la cuestión de la coinstanciación de las propiedades cognitivas y las computacionales (que son los tipos de propiedades a las que se aluden en el nivel de procesamiento de la información) más que en la cuestión de la implementación física. En este sentido, se requiere explicar por qué un determinado estado cognitivo es condición suficiente para instanciar un determinado estado computacional y otro estado computacional es condición

---

<sup>10</sup> El ejemplo es de Cummins (1989), p. 89.

suficiente para dar lugar a otro estado intencional.<sup>11</sup> Dado que la teoría representacional de la mente tiene una respuesta a esta cuestión dejaré su desarrollo para la segunda parte de este capítulo. En lo que sigue, voy a dar por sentado que una función cognitiva se instancia en la mente/cerebro a través de las funciones de interpretación y realización.

### 1.5. La definición

La última cuestión se relaciona con aspectos en torno a la definición de representación. El problema principal consiste en responder a la pregunta: qué es para algo representar. Pero también hay aspectos relacionados con esta cuestión como, por ejemplo, los distintos tipos de representaciones: explícita, tácita, implícita (Dennett 1987, Cummins 1986); locales y distribuidas (Hinton, McClelland y Rumelhart 1986; van Gelder 1991); indicativas e imperativas (Millikan 1986), y la cuestión de la relación representacional.

Una definición de mínima de representación afirma que algo representa otra cosa si sustituye a o está por (*stand in for*) otra cosa, que en general se suele sostener que es algún rasgo del mundo (Haugeland 1991). Además, se suele aceptar que las representaciones tienen propiedades semánticas/intencionales (por ejemplo, significado, referencia, condiciones de verdad, coherencia inferencial). Así, hay que ofrecer alguna teoría acerca de cómo las representaciones adquieren sus propiedades semánticas/intencionales, esto es, hay que explicar la relación representacional entre el contenido de los estados intencionales y los objetos en el mundo, y para ello se postulan semánticas del contenido mental. Estas semánticas al proponer el tipo de relación que mantiene una representación con lo representado, especifican mucho más la definición de representación al tiempo que

---

<sup>11</sup> Cf. Fodor (1994), p. 10. Por razones de simplificación no tomo en cuenta la cuestión de la realizabilidad múltiple.

intentan dar cuenta de qué es para un estado mental tener un contenido específico.<sup>12</sup>

En general, todas las semánticas propuestas están pensadas para estados intencionales y pueden agruparse en cuatro tipos dependiendo de la relación representacional: las semánticas de rol conceptual (Block 1986, Harman 1982), la semántica interpretacional (Cummins 1989), las teorías covariacionales (Dretske 1981, Fodor 1990b), y las teorías teleológicas (Millikan 1984, Papineau 1987, McGinn 1989). La semántica del rol conceptual sostiene que el contenido de las representaciones está determinado por el rol funcional / inferencial / causal / computacional que posee una determinada representación en la economía cognitiva del individuo. Así, quienes sostienen la versión inferencial afirman que el contenido depende de las conexiones inferenciales en las que esté involucrada esa representación o si se apunta al rol causal, se dirá que el contenido depende de las relaciones causales de la representación con los inputs, los outputs y otras representaciones en el sistema cognitivo.<sup>13</sup>

La semántica interpretacional es la única propuesta actualmente disponible en la filosofía de la psicología que está pensada para estados representacionales no intencionales. Las representaciones se definen a partir de la relación inducida por la proyección interpretativa entre los elementos de la función física y la función cognitiva instanciada. Así, la noción de representación tiene sentido porque hay simulación (la instanciación es un caso especial porque relaciona objetos abstractos y sus instancias físicas) e interpretación (la proyección de procesos computacionales sobre otros procesos). Cummins (1989, p. 129) define representación de la siguiente manera: “[una representación] *R* representa *x* en [el sistema computacional] *S* sii hay funciones *g*, *f*, y una Interpretación *I* tal que *S*

---

<sup>12</sup> Generalmente, se suele desplegar el problema del contenido en dos aspectos: en qué consiste que un estado de un sistema sea un estado intencional y qué hace que un estado tenga un contenido específico. Casi todos los filósofos que propusieron respuestas para lo segundo son funcionalistas con respecto a lo primero. Así, que un estado físico de un sistema tenga un contenido intencional depende de su rol funcional en el sistema, esto es, de las conexiones con los inputs, los outputs y otros estados funcionales. Sin embargo, el contenido del estado intencional no tiene por qué individuarse en función de su rol funcional. Es de la cuestión del contenido de la que me ocuparé en los capítulos siguientes.

<sup>13</sup> En el capítulo 5 digo algo más de la semántica del rol conceptual.

satisface  $g$ ,  $g$  simula  $f$  bajo  $I$  e,  $I(R) = x$ ." Así, los objetos de la computación vistos bajo interpretación son representaciones.

En las teorías covariacionales la noción básica es la de información que se define en términos de relaciones nómicas entre eventos. Los agentes cognitivos son vistos como sistemas que procesan información por medio de mecanismos de nivel inferior puramente físicos. El contenido mental queda fijado por una relación de covariación causal entre algo en el mundo y una representación que de esta manera porta información acerca de eso en el mundo. Las teorías teleológicas sostienen que los contenidos se determinan por las condiciones que permiten que un organismo cumpla con su función biológica. Estas teorías dan cuenta de las funciones en términos evolutivos apelando a la selección natural que actúa sobre las especies para producir adaptaciones al entorno.

En los capítulos 3 y 4 me ocuparé de estos dos últimos tipos de semánticas del contenido intencional.<sup>14</sup> Voy a centrar mi crítica a las semánticas del contenido intencional en dos de las propuestas más desarrolladas en el marco de las teorías covariacionales y las teleológicas: la Semántica Informacional de J. Fodor y la Teleosemántica de R. Millikan. En el capítulo 3 expondré ambas teorías y distinguiré las tesis básicas que comparten y en el capítulo 4 presentaré una propuesta naturalista alternativa. Mi objetivo principal será mostrar que el proyecto de una semántica para estados intencionales es superfluo. En lo que sigue me ocuparé de exponer el marco general en el que las semánticas del contenido pretenden insertarse, esto es, la TRM, y veremos cómo ésta responde a las cinco cuestiones planteadas en esta primera parte.

## 2. La Teoría Representacional de la Mente

---

<sup>14</sup> Dejo a un lado la semántica del rol conceptual porque tiene más bien la forma de un proyecto general y no parece haber ninguna teoría lo suficientemente desarrollada y específica a la manera de las teorías covariacionales o teleológicas. Por otro lado, dado que la semántica interpretacional no está concebida para el contenido intencional en particular, sino para todo estado representacional

La TRM surge en el marco de la filosofía de la psicología para dar cuenta de la naturaleza del pensamiento y del pensar.<sup>15</sup> Como hemos visto en el capítulo anterior, la variedad de pensamientos de los cuales se han ocupado los filósofos y psicólogos de la mente es de las APs, de manera que la TRM está pensada particularmente para los estados mentales intencionales. Como bien dice Fodor, la TRM “es una vaga confederación de tesis” (1998, p. 24). Pero además, cuando se suele reconstruirlas se lo hace de manera tendenciosa dependiendo de quién es el filósofo que las reconstruye, punto que se aplica en particular a Fodor, quien no sólo fue el precursor de esta teoría sino que fue el que formuló sus tesis de una manera explícita. Intentaré reconstruir las tesis principales de manera de ir de las compartidas por cualquiera que adhiera a la TRM a las más controvertidas, dejando de lado algunas que claramente no forman parte pero que a veces se las incluye en algunas formulaciones de la TRM.<sup>16</sup>

## 2.1. La teoría

Se puede reconstruir las tesis de la TRM de la siguiente manera:

### 1. *Naturaleza de los estados mentales.*

Para cualquier sujeto  $S$  y cualquier actitud proposicional  $A$  que se tenga hacia la proposición  $P$ , hay una única y distintiva relación psicológica  $R$  y una representación mental  $RM$  tal que ‘ $SA$  que  $P$  si y sólo si:  $S$  mantiene  $R$  con  $RM$  y  $RM$  significa que  $P$ . (Fodor 1987, p. 38; Field 1978, p. 37)

---

no considero que está sujeta a las críticas que desarrollaré en el capítulo 4. Es más, a lo largo de la tesis, quedará claro que adhiero a alguna forma de semántica interpretacional.

<sup>15</sup> A veces se utiliza indistintamente las denominaciones TRM y TRCM y así lo haré también, pero hay que tener presente que para que la TRM sea una TRCM hay que adherir a la tesis de que los procesos mentales son computacionales en el sentido en que la TRCM los especifica. Dado que los psicólogos y filósofos que aceptan el nivel de procesamiento de la información suelen aceptar que el procesamiento es computacional, en ese sentido son asimilables.

## 2. *Naturaleza de los procesos mentales.*

Los procesos mentales consisten en secuencias causales de instancias de representaciones mentales. (Fodor 1987, p. 39)

Cualquiera que adhiera a la TRM tiene que sostener de mínima ambas tesis. La tesis 1 está compuesta de dos subtesis:

1.1. Los estados intencionales son estados relacionales (*vs.* estados monádicos)

1.2. La relación *R* se establece entre un sujeto (u organismo) y representaciones mentales (*vs.* que el objeto de las actitudes sean las emisiones lingüísticas o las proposiciones).

Lo que le da el carácter representacional a esta teoría de la mente es, en particular, la subtesis 1.2. Actualmente, la mayoría de los filósofos de la mente aceptan el enfoque relacional, sin embargo no todos los que aceptan el enfoque relacional consideran que los objetos de las actitudes son las representaciones mentales y en ese sentido, sólo los que adhieren a 1.2 sostienen una TRM. Voy a desarrollar 1.1 y 1.2, luego introduciré dos subtesis acerca de las representaciones mentales, la 1.3 que es acerca de la instanciación de las representaciones mentales y que es aceptada por todos los que adhieren a la TRM y la 1.4 que es acerca de la naturaleza de las representaciones mentales, la hipótesis del lenguaje del pensamiento, la cual no es aceptada por todo aquel que adhiere a la TRM.

Tanto para el enfoque relacional como para la idea de que el objeto de las actitudes son las representaciones mentales hay argumentos conceptuales y empíricos. El argumento empírico que subyace a ambas tesis y al resto de la tesis comprendidas en la TRM (incluso las que aún no hemos mencionado) está basado en que las teorías psicológicas suponen la TRM, de manera que si nuestras mejores teorías psicológicas acerca de las capacidades cognitivas superiores (fijación de creencias perceptivas, aprendizaje de conceptos y toma de decisiones) son correctas

---

<sup>16</sup> Por ejemplo, Fodor (1998, cap. 1) incluye tesis acerca de la semántica de las representaciones mentales como parte de su propia concepción de la TRM. Sin embargo, otros defensores de la TRM sostienen otros tipos de semánticas.

entonces, la TRM es cierta (Fodor 1975, 1987). Veamos 1.1 y algunos de los argumentos conceptuales en su favor.

Los estados intencionales son estados relacionales en el sentido en que los verbos proposicionales que denotan actitudes (como creer, desear, esperar, y similares) son predicados relacionales que expresan relaciones entre un individuo y el objeto de la actitud, por ejemplo en 'Alex cree que el pasto es verde', Alex mantiene una relación de creencia con el objeto de la creencia expresado en la cláusula-que. El enfoque relacional se opone al enfoque monádico (por ejemplo, Quine 1956) que considera que las APs no son relaciones sino propiedades de un individuo (de manera que se prescinde del objeto de la actitud). Así, 'Alex cree que el pasto es verde' contiene una expresión referencial ('Alex') y un predicado monádico que no tiene estructura lógica interna, es atómico ('cree-que-el-pasto-es-verde').

Algunos de los argumentos conceptuales en favor del enfoque relacional son los siguientes (Fodor 1978):

a- *argumento del aprendizaje*: si los predicados que figuran en las APs fueran atómicos, sería imposible aprender un lenguaje (habría que adquirir una lista infinita de predicados).

b- *argumento del contenido en común*: por un lado, un mismo contenido puede acompañar distintas actitudes de manera que se puede decir que éstas tienen algo en común (desde un enfoque no relacional este hecho sería un mero accidente), y por el otro, se puede dar cuenta de un desacuerdo de creencias porque son acerca del mismo objeto, individuo, etc., es decir, tienen algo en común (desde un enfoque relacional las creencias no tendrían nada en común).

c- *argumento de las condiciones de verdad*: según el enfoque no-relacional sería un mero accidente que las condiciones de verdad de lo que cree Alex en 'Alex cree que el pasto es verde' sean las mismas que las de la expresión 'el pasto es verde'.

Con respecto a *R*, las distintas relaciones que un individuo mantiene con las representaciones mentales constituyen los distintos tipos de actitudes: creencia, deseo, y similares. Las relaciones son entendidas en términos funcionales, por



ejemplo lo que hace que un estado mental en particular sea una creencia y no un deseo es su rol funcional (causal o computacional), esto es, que suele ser el output de mecanismos perceptivos e input para sistemas de inferencia en los que interactúa con otros estados mentales para producir comportamiento u otros estados mentales. Consideradas en términos del rol computacional, cada actitud se individúa en función de determinadas operaciones que se aplican a las representaciones. Se suele usar la expresión 'caja de creencias' (o de deseos, intenciones, y demás, Schiffer 1981) para cubrir cualquier tipo de rol que se crea que tiene cada tipo de actitud, la imagen que se desprende de esta metáfora es la de representaciones mentales que van a la caja de creencias o deseos según el tipo de operaciones que se les aplique o el rol causal que tengan. Así, creer que tal y cual es tener en la caja de creencias una instancia de una representación mental que significa tal y cual.

La subtesis 1.2 afirma que las APs son relaciones entre un organismo y representaciones mentales. Los objetos de las actitudes no son proposiciones (entendidas como sentidos fregeanos o conjuntos de mundos posibles, Stalnaker 1976) ni oraciones de un lenguaje público (Carnap 1947, Davidson 1968) sino entidades en la mente. Tal como hemos visto, las representaciones mentales tienen dos propiedades fundamentales, son semánticamente evaluables y tienen un rol causal en la producción de la conducta. Para que esto último sea posible se requiere sostener la siguiente tesis:

### 1.3. Las representaciones mentales se instancian en estructuras físicas en el cerebro.

La única manera de que las representaciones mentales sean cognitivamente relevantes para explicar la conducta es que formen parte de la estructura causal del sistema. Una instancia de una representación mental es un objeto físico con propiedades semánticas. Se apela a la relación de instanciación, realización o implementación para dar cuenta de la conexión entre lo mental y lo físico (o lo computacional).

Si bien las representaciones comparten con las proposiciones y las emisiones lingüísticas la propiedad de ser semánticamente evaluables, no ocurre lo mismo con la propiedad de tener un rol causal. Las proposiciones son entidades abstractas y en tanto tales no tienen un rol causal, además en tanto que no están instanciadas en la cabeza de los individuos, surgirían los siguientes problemas: ¿cómo podrían diferir cognitivamente dos mentes?, ¿cuál sería la relación entre las mentes y las entidades abstractas que no fuera la poco explicativa y muy problemática de 'captación'? Las proposiciones sólo podrían tener un rol causal si estuvieran instanciadas en la mente/cerebro, la postulación de representaciones internas que expresan proposiciones es la manera de dar cuenta de la relación entre un organismo y una proposición que es el contenido de las APs (Fodor 1978).

A diferencia de las proposiciones, las emisiones lingüísticas son entidades concretas pero que no están en las cabezas de los individuos. Aquellos que sostienen este enfoque suelen considerar que las creencias son disposiciones a respuestas verbales. Pero, las representaciones mentales no pueden ser disposiciones. Mientras que una entidad en la mente instanciada en objetos físicos estaría en condiciones, en principio, de tener rol causal, las disposiciones no, porque no puede haber causación disposicional sin que haya causación no-disposicional. Dada una propiedad disposicional como por ejemplo la fragilidad, se requiere que algo pase para que se manifieste dicha disposición. Se supone que la fragilidad causaría la rotura de un jarrón al tirarlo, pero se requiere de un evento que cause que la disposición se manifieste en la rotura del jarrón (por ejemplo, tirar el jarrón), de manera que, entre una disposición y otra tiene que haber un evento causal no-disposicional. Puesto que según la concepción disposicional, lo mental estaría constituido por puras disposiciones, y dado que una cadena de puras disposiciones no pueden causarse entre sí, entonces no sería posible la causación mental.

Además, si las emisiones lingüísticas fueran el objeto de las APs, habría otros problemas. Por ejemplo, no se podrían tener creencias si no se tiene un lenguaje natural, no se explicaría el hecho de que hay creencias inexpresables, y no se podría dar cuenta de la adquisición del lenguaje porque dado que el aprendizaje

de un lenguaje natural consiste en la confirmación de hipótesis y este proceso supone la participación de estados mentales como creencias, expectativas, y similares, y dado que tener estados mentales supone ya tener un lenguaje natural, estaríamos en un círculo explicativo vicioso.

El problema de considerar el objeto de las APs como oraciones del lenguaje público puede solucionarse, según Fodor (1975), si en su lugar se piensa que las representaciones mentales son oraciones de un lenguaje interno, un lenguaje del pensamiento (o mentalés). Algunos defensores de la TRM adhieren a 1.4:<sup>17</sup>

1.4. Las representaciones mentales tienen una estructura simbólica específica: una semántica composicional y una sintaxis combinatoria. (Fodor y Pylyshyn 1988)

Las representaciones mentales son oraciones de un lenguaje interno que no es ninguno de los lenguajes naturales pero tiene las mismas propiedades semánticas y sintácticas que los lenguajes naturales. La sintaxis combinatoria consiste en que las representaciones moleculares tienen representaciones atómicas que respetan la relación parte-todo. Una representación mental compleja, por ejemplo 'Mariana ama a Ariel' está constituida por los constituyentes sintácticos: 'Mariana', 'amar', y 'Ariel' que mantienen entre sí ciertas relaciones sintácticas o estructurales. La semántica composicional consiste en que el contenido de la representación molecular depende del contenido de los constituyentes sintácticos y sus relaciones estructurales. Los contenidos de los constituyentes son independientes del contexto, es decir que, dado un constituyente, éste realiza siempre la misma contribución semántica a la oración en la que figura.

La hipótesis del lenguaje del pensamiento pretende ser una tesis empírica y su postulación se basa tanto en razones empíricas como conceptuales. Los

---

<sup>17</sup> Algunos de los defensores de la TRM que adhieren a 1.4 son Block (1986), Rey (1991), Lycan (1993) y Field (1978). Los que la rechazan explícitamente son, entre otros, Lewis (1972), y Barwise y Perry (1983). Los que se mantienen neutral son, entre otros, Loar (1983), Dretske (1981) y Armstrong (1980). Algunos consideran que el lenguaje interno es un lenguaje natural (Harman, 1982), otros se mantienen neutral con respecto a si el lenguaje del pensamiento no es ninguno de los lenguajes naturales o es alguno de ellos (Botterill y Carruthers 1999).

argumentos principales para postular un lenguaje del pensamiento son los siguientes:

a. *argumento de los procesos psicológicos* (Fodor 1987, pp. 202-7): las teorías psicológicas dan cuenta de los procesos mentales en términos computacionales, los procesos computacionales operan sobre un medio representacional y sólo son sensibles a la estructura constitutiva (sintaxis) de las representaciones, por ende las representaciones son objetos estructurados sintácticamente.

b. *argumento de la productividad del pensamiento* (Fodor 1987, pp. 207-16; Fodor y Pylyshyn 1988): es un hecho empírico que tenemos la capacidad de tener pensamientos infinitos (competencia) aun cuando sólo podamos pensar un número limitado (realización). Dado que somos seres finitos sólo podemos tener recursos finitos, por ende la productividad sólo es posible gracias a la recursividad que permite la sintaxis combinatoria del lenguaje del pensamiento.

c. *argumento de la sistematicidad y composicionalidad del pensamiento* (Fodor 1987, pp. 207-16; Fodor y Pylyshyn 1988; Fodor y McLaughlin 1990): es un hecho empírico que la habilidad para tener pensamientos está conectada con la habilidad para tener otros pensamientos (un organismo que puede pensar aRb, puede pensar bRa y viceversa). Se pueden producir pensamientos con combinaciones distintas de los mismos constituyentes sintácticos y entenderlos puesto que el significado de la expresión compleja depende de los significados independientes del contexto de cada constituyente sintáctico y de las relaciones sintácticas entre los constituyentes. Así, la sistematicidad que es una propiedad sintáctica del pensamiento se explica en función de una propiedad semántica que es la composicionalidad.<sup>18</sup>

Pasamos ahora a la tesis 2 acerca de los procesos mentales. En general, los que sostienen una TRM adhieren a un enfoque computacional de los procesos

---

<sup>18</sup> Tanto la productividad (la habilidad para producir y comprender infinitas oraciones) como la sistematicidad (la habilidad para producir y entender algunas oraciones está conectada de manera intrínseca con la habilidad para producir y entender otras oraciones) son propiedades de la capacidad lingüística. Se asume que lo son también del pensamiento dado que la función del lenguaje es expresar el pensamiento (Fodor 1987, apéndice).

mentales. Los que adhieren al enfoque computacional clásico (Turing) sostienen una TRCM. En este sentido, 2 se subdivide en las siguientes tesis:

2.1. Los procesos mentales consisten en computaciones sobre representaciones mentales. (Fodor 1998)

2.2. Los procesos mentales son sensibles a la forma (sintaxis) de las representaciones mentales. ('condición de formalidad', Fodor 1980)

Las computaciones son procesos causales que operan transformaciones entre representaciones mentales respetando de manera fiable las propiedades semánticas de las mismas (tesis 2.1). Una representación mental desde un punto de vista semántico tiene un significado expresado en su contenido y desde el punto de vista sintáctico se presenta de cierta forma. Por ejemplo, las representaciones mentales AGUA y H<sub>2</sub>O tienen el mismo contenido porque significan lo mismo pero distintos modos de presentación porque las formas de las expresiones son distintas (tienen distinta estructura, en este caso, distintas combinaciones de formas de letras). Así, los modos de presentación son los vehículos sintácticos de los contenidos semánticos (ambos, vehículo sintáctico y contenido semántico conforman las representaciones mentales).<sup>19</sup>

Los procesos computacionales operan en virtud de la sintaxis de las representaciones mentales (tesis 2.2). Esto significa que sólo pueden afectar un sistema físico en virtud de su forma (la sintaxis es una propiedad física de orden superior, una propiedad abstracta de la forma de un símbolo, cf. Fodor 1987, p. 40). Así, dos pensamientos que poseen el mismo contenido son individuados en virtud de que forman diferentes cadenas de símbolos (son ocurrencias de estados mentales-tipo diferentes). Su rol causal (/inferencial) no es el mismo y esto sólo parece poder explicarse en virtud de que sus formas (o sintaxis o modos de

---

<sup>19</sup> De manera que, más específicamente, estar en un estado mental consiste en una relación de tres lugares entre un organismo, una proposición (el contenido) y un modo de presentación. Cf. Fodor (1994), p. 47.

presentación) no son las mismas.<sup>20</sup> En tanto que la sintaxis codifica aspectos semánticos que determinan los roles inferenciales de los símbolos, los procesos computacionales respetan la coherencia semántica del pensamiento. Así, son posibles mecánicamente razonamientos semánticamente coherentes (y que preservan la verdad).

Cuando Fodor se refiere a la relación entre la sintaxis y la semántica de las representaciones mentales, utiliza expresiones del tipo: la sintaxis “respeto” la semántica (1987, pp. 32-33), hay un “paralelismo” entre sintaxis y semántica (1987, p. 34), las relaciones semánticas entre los símbolos pueden ser “imitadas” por sus relaciones sintácticas (1987, p. 41). La idea es la siguiente:

Con todo, la máquina está diseñada de tal manera que transformará un símbolo en otro si y sólo si las proposiciones expresadas por los símbolos que se transforman de esa manera guardan ciertas relaciones *semánticas*; por ejemplo, la relación que mantienen las premisas con la conclusión en un argumento válido. Semejantes máquinas -los ordenadores, por supuesto- son precisamente entornos en los que la sintaxis de un símbolo determina su papel causal de una manera que respeta su contenido. Esta es, según pienso, una idea estupenda, sobre todo porque funciona. (Fodor 1987, p. 41)

La sintaxis permite conectar la causalidad con la semántica: la causalidad es sensible a la sintaxis, la sintaxis es sensible a la semántica y por ende, la causalidad es sensible a la semántica. El argumento principal para adherir al enfoque computacional de los procesos mentales es que permite dar cuenta de un hecho empírico acerca del proceso del pensamiento que es el paralelismo entre las relaciones causales y las relaciones semánticas entre nuestros pensamientos o dicho de otra manera, la sincronización del rol causal de una representación con su contenido (Fodor 1987, pp. 39-42).

---

<sup>20</sup> Así, pensamientos con la misma estructura juegan el mismo rol inferencial. Por ejemplo, las representaciones mentales de estructura conjuntiva pueden ser sujeto del mismo rol inferencial de

## 2.2. La teoría y el problema de las representaciones

Dadas las tesis principales de la TRM, veamos cómo responde a las cinco cuestiones acerca de las representaciones mentales desarrolladas en la primera parte.

En primer lugar, los *criterios para postular representaciones* responden a motivaciones similares a las que hemos visto en la primera parte del capítulo, sólo que enfocadas en la explicación de la acción intencional por medio de las APs. En este caso, una de las formulaciones del argumento a la mejor explicación sería la siguiente: si no se postularan representaciones mentales no se podría dar cuenta de cómo los estados intencionales tienen poder causal en la producción de la conducta.

Se reconoce que pueden darse casos de actitudes sin representaciones explícitas (creencias disposicionales) o representaciones explícitas sin actitudes (en el nivel subpersonal hay representaciones instanciadas físicamente pero dado que las personas no tienen acceso a sus contenidos no pueden tener ninguna relación epistémica, es decir, ninguna actitud, hacia ellas), pero se asume que los casos nucleares de la TRM son casos de APs y se sostiene que todo caso de AP va acompañado siempre de representaciones mentales. En otras palabras, si instancias de actitudes participan de procesos mentales y los procesos mentales operan sobre representaciones, entonces, las instancias de actitudes se tienen que corresponder con instancias de representaciones (Fodor 1987, cap. 1). En este sentido, la TRM es una propuesta para la explicación de los estados intencionales del nivel personal.

La TRM es una teoría acerca de una parte de la cognición, aquella relacionada con el pensamiento y lo que afirma es que siempre que hay una actitud (un estado mental intencional) hay representaciones y que éstas se requieren para dar cuenta de las actitudes, de manera que la noción primitiva que explica la intencionalidad es la de representación. Cuando se da cuenta de la naturaleza de las representaciones se lo hace pensando en las propiedades que deben tener si

---

simplificación de la conjunción.

participan de estados mentales intencionales. Dado que hay estados mentales que no son intencionales pero de los que igualmente participan representaciones, la TRM no dice nada acerca de por qué es necesario postular representaciones para esos estados del nivel subpersonal. Puesto que la TRM no dice nada acerca de por qué es necesario postular representaciones para dar cuenta de las capacidades cognitivas que no requieren de actitudes, más allá de que esté en lo correcto o no acerca de aquellas que requieren APs, la TRM no puede ser el relato completo acerca de las representaciones mentales.

Ante esta situación, se podría pensar que la TRM simplemente recorta un dominio de lo cognitivo: el representacional intencional. Sin embargo, hay razones de peso para pensar que la TRM asimila sin más el ámbito de lo representacional al ámbito de lo intencional. Esto se verá a lo largo de los capítulos restantes pero, en principio, podemos ver que es así cuando revisamos los *requisitos para una teoría de la representación*. La TRM asume que el requisito fundamental de una teoría representacional es explicar el fenómeno de la intencionalidad ((d) requisito del contenido intencional). La TRM da la base indispensable para comprender la naturaleza de los estados intencionales que debe completarse con una semántica del contenido que explique cómo las representaciones mentales adquieren su contenido o cómo se explica la noción de representación en términos naturalistas ((c) requisito naturalista), y cómo puede darse el fenómeno de la representación errónea ((e) requisito de la representación errónea), de manera de tener la explicación completa de la conducta de los individuos en función de sus estados intencionales.

Así, según la TRM el rol explicativo de las representaciones mentales es fundamental en las explicaciones intencionales del comportamiento humano. Y, tal como venimos viendo, se cree que el objetivo de la psicología es dar cuenta de la conducta en función de apelar a estados intencionales. La TRM adhiere a la idea de que la teoría debe sustentar, o al menos ser compatible con, el rol explicativo que las ciencias cognitivas le otorgan a las representaciones ((a) requisito explicativo), pero se equivoca con respecto al rol que las teorías psicológicas le otorgan a las representaciones mentales. En el capítulo 5 quedará claro que el objetivo de la



psicología cognitiva no es dar cuenta de la totalidad de la complejidad de la acción intencional.

En cuanto a (f) el requisito de la integración cognitiva, la TRM debe complementarse con una teoría acerca de la arquitectura mental que permita dar cuenta de cómo las representaciones se integran con otras representaciones. Por ejemplo, cómo las representaciones que participan en los estados intencionales se integran con las de los estados perceptivos, motivacionales, de atención, etc. y con representaciones almacenadas. En general, los defensores de la TRM adhieren a una arquitectura cognitiva clásica, que tiene innumerables problemas para dar cuenta de la integración, más alguna teoría acerca de la distinción entre sistemas modulares y sistemas centrales.<sup>21</sup>

En relación a (b) el requisito implementacional, la TRM acata la idea de que la teoría debe ser compatible con las mejores explicaciones científicas acerca de qué entidades cumplen el rol representacional en la mente/cerebro. Dado que este requisito se relaciona con dos de las cuestiones acerca de las representaciones que nos falta ver, pasaré directamente a ellas. La cuestión acerca del *formato representacional* adquiere distintas respuestas. La mayoría de los que adhieren a la TRM consideran que el formato es simbólico, pero tal como hemos visto en la primera parte también hay quienes sostienen que el formato es pictórico (Kosslyn y Pomerantz 1977) o simbólico pero sin alguna de las propiedades que la TRCM le atribuye a las representaciones (por ejemplo, algunos conexionistas consideran que las representaciones no tienen la estructura sintáctica de los constituyentes clásicos (Smolensky 1991, 1994) y por ende, el tipo de computación no es la clásica. De todas formas estas teorías son computacionales y representacionistas, cf. Haugeland 1991, Clark y Toribio 1994).

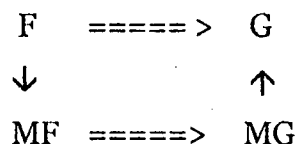
La TRCM asume que los estados intencionales se implementan en estados computacionales, así *la implementación de las representaciones* se da en objetos físicos del sistema. Si se toma en cuenta los niveles de explicación de una teoría

---

<sup>21</sup> En realidad, la TRCM es una teoría cognitiva clásica (esto se puede ver en Fodor y Pylyshyn 1988). En relación a los problemas del cognitivismo clásico para ofrecer una visión integrada de la cognición, véase, entre otros, Dreyfus (1979) y Clark (1989). En relación a la distinción entre sistemas centrales y módulos, véase Fodor (1983).

cognitiva: el nivel computacional, algorítmico e implementacional, la TRCM intenta dar una explicación en los dos primeros niveles, el tercero sería el de la implementación física al que, siguiendo los postulados funcionalistas, no se le otorga un papel de peso. Así, Fodor (1994) propone una respuesta al problema de la coinstanciación de las propiedades computacionales e intencionales, en términos de leyes intencionales. Son las leyes intencionales (cuya forma es: 'los eventos Fs causan eventos Gs, *ceteris paribus*) las que están implementadas en mecanismos computacionales.

Una teoría de la implementación responde a la pregunta: ¿Cómo los Fs causan los Gs? y la respuesta consiste en especificar un mecanismo (computacional) en el que la instanciación de F es suficiente para la instanciación de G. La figura que expresa esta relación es la siguiente (las flechas dobles indican que el antecedente es condición *causalmente* suficiente para el consecuente y las flechas simples indican que el antecedente es condición suficiente para la *instanciación* del consecuente):<sup>22</sup>



El hecho de que los MFs causan los MGs en el nivel computacional explica el hecho de que los Fs causen los Gs en el nivel intencional. Pero, por supuesto se necesita una explicación de por qué F es condición suficiente para instanciar MF y MG es condición suficiente para dar lugar a G (Fodor 1994, p. 10). La respuesta a esta cuestión de coinstanciación de propiedades computacionales e intencionales, en el fondo sigue siendo la misma respuesta que relacionaba a la sintaxis con la semántica de los símbolos. Antes vimos que la sintaxis "respetar" la semántica. Aquí, "respetar" significa que la relación entre el mecanismo computacional y los contenidos intencionales es metafísicamente contingente, su coinstanciación depende de hechos acerca del mundo.

<sup>22</sup> Según Fodor (1994, p. 10), por razones de simplicidad el diagrama supone que  $F \Rightarrow G$  no está múltiplemente realizado. Los distintos tipos de flechas es un agregado a la figura.

Esto no significa decir que es accidental sino que los hechos se obtienen tanto en el mundo efectivo como en todos aquellos mundos donde las leyes psicológicas se implementan de la misma forma que en el mundo efectivo.<sup>23</sup> Sin embargo, este intento de dar cuenta de la coinstanciación de propiedades intencionales y computacionales depende completamente de la postulación de leyes intencionales. Muchos filósofos creen que la psicología cognitiva científica es un asunto acerca de establecer leyes intencionales (causales).<sup>24</sup> En el capítulo 4 veremos que la psicología cognitiva no requiere de este tipo de leyes y en algún sentido, de ningún tipo de leyes especiales.

Nos queda por ver el último problema en torno a las representaciones que son cuestiones relacionadas con la *definición de representación*, en particular la cuestión acerca de la relación de representar o la semántica del contenido. La TRM es neutral acerca de cómo las representaciones mentales adquieren su contenido. Algunos defensores de la TRM sostienen una teoría covariacional (Fodor), otros una semántica del rol conceptual (Block), otros una semántica interpretacional (Cummins) y otros semánticas teleológicas (Millikan). Todas estas semánticas, en tanto son externalistas (o postulan algún componente externalista), tienen problemas de compatibilización con el enfoque computacional porque tienen que explicar cómo es que supervienen propiedades extrínsecas (las intencionales) a propiedades intrínsecas (las computacionales).<sup>25</sup>

Sin embargo, no es de este problema del cual me ocuparé sino de la viabilidad del proyecto de una semántica para las representaciones mentales. En el capítulo siguiente voy a reconstruir las tesis principales de lo que denomino el

---

<sup>23</sup> Habría dificultades para dar cuenta de la coinstanciación de las propiedades computacionales e intencionales. Fodor (1994) intenta dar una solución a los casos Frege y los casos Gemelos que serían contraejemplos para esta coinstanciación. Los casos Frege muestran que la identidad de estados intencionales no es suficiente para la identidad de la implementación (porque más de un mecanismo implementacional puede corresponderse con el mismo estado intencional) y los casos Gemelos muestran que la identidad de implementación no puede ser suficiente para la identidad del estado intencional (porque más de un estado intencional puede corresponderse con el mismo mecanismo implementacional).

<sup>24</sup> El calificativo 'muchos' realmente abarca una lista enorme de filósofos de la psicología y de la mente, por eso me resulta imposible enumerarlos.

<sup>25</sup> Las teorías de rol conceptual solipsistas (aquellas que no tienen un componente externalista) también tienen problemas de compatibilización con el enfoque computacional pero son de otro tipo. Cf. Cummins (1989), cap. 9.

proyecto naturalista científicista de una semántica del contenido, a partir de la semántica informacional de J. Fodor y la teleosemántica de R. Millikan que se inscriben respectivamente en los enfoques covariacionales y teleológicos. Estas teorías responden a lo que considero un proyecto que no sólo no forma parte de la psicología cognitiva, como es de esperar, sino que ésta no lo requiere. Mi intención es por un lado, mostrar que para dar cuenta del fenómeno de la intencionalidad no se requiere del desarrollo de semánticas del contenido (capítulo 4), y por el otro, que la psicología cognitiva tampoco requiere de semánticas del contenido (capítulo 5).

## 3

## La semántica naturalista científicista

En los capítulos anteriores hemos visto las tesis básicas de la psicología y de la TRM. También hemos visto las respuestas de la TRM a ciertas cuestiones básicas en torno a las representaciones mentales. En el desarrollo de estos temas se ha advertido cómo la TRM es un intento por introducir cuestiones filosóficas acerca de la intencionalidad en el ámbito de la psicología. Asimismo se ha señalado que para una explicación adecuada del fenómeno de la intencionalidad, tal como lo entienden los filósofos de la psicología, no basta con esta teoría sino que se requiere del desarrollo de una semántica intencional. La filosofía de la psicología considera que las representaciones mentales postuladas por la psicología requieren de una semántica que de cuenta de cómo las representaciones son *acerca de* algo en el mundo externo al individuo.

En este capítulo me ocupo de las dos semánticas más desarrolladas a las que agrupo bajo la denominación de *Proyecto Naturalista Científicista* (PNC). Estas semánticas son la *Semántica Informativa* de Jerry Fodor y la *Teleosemántica* de Ruth Millikan. Las he elegido porque creo que se ganará en profundidad con respecto a la comprensión de en qué consiste un proyecto así si se desarrollan las dos teorías más elaboradas que disponemos actualmente. Ambas teorías compiten por la explicación del fenómeno de la intencionalidad de las representaciones mentales dentro del marco de la TRM,<sup>1</sup> a la vez que representan las dos tendencias

---

<sup>1</sup> Habría razones para pensar que la teleosemántica no intenta ser un complemento de la TRCM (si bien sí de una TRM que sea computacional al estilo conexionista, ver nota 27). Sin embargo, se ha intentado compatibilizar esta semántica con la TRCM (así como también la semántica informativa y la del rol conceptual). Por ejemplo, Cummins (1989) es una crítica a estas tres semánticas y una defensa de su semántica interpretacional con respecto a la cuestión de cuál es la semántica apropiada para la TRCM. Lo que esto muestra es que los filósofos consideran a estos cuatro tipos de semánticas como propuestas que intentan dar cuenta del contenido de los estados computacionales. Es en este sentido que considero a la teleosemántica y la semántica informativa

más significativas dentro del proyecto científicista de naturalización de los contenidos mentales.

El calificativo *naturalista* alude a uno de los requisitos que los filósofos suelen considerar que una teoría de las representaciones debe satisfacer. El requisito naturalista (requisito (c) del §1.2 del capítulo anterior) afirma que la teoría no debe apelar a términos semánticos/intencionales para explicar la noción de representación. Si de lo que se trata es de dar cuenta de las propiedades intencionales de las representaciones, la explicación no debe apelar a las mismas puesto que se caería en un círculo. Este círculo podría ser iluminador, en el sentido de que las nociones intencionales como la de representación se explicarían en función de otras nociones intencionales (como “creencia”, “significado”, y similares), lo que daría cuenta de un rasgo peculiar de la intencionalidad. Pero, para la mayoría de los filósofos de la mente, sería un círculo vicioso dado que tendría como consecuencia ontológica la no pertenencia de las propiedades intencionales al mundo físico.

El PNC es un proyecto de naturalización del contenido mental que se toma muy en serio su tarea redentora de las propiedades intencionales. Por “propiedades intencionales” se entiende una subclase de las propiedades mentales que consiste en el estar (de un organismo) en un determinado estado de creencia, deseo, intención, y similares, que es acerca de algo en el mundo.<sup>2</sup> Es una tarea redentora porque la motivación básica que guía a toda empresa de naturalización del contenido consiste en la creencia de que si no se vinculan las propiedades intencionales con algo del mundo natural, éstas quedan eliminadas de la ontología

---

como las dos alternativas más desarrolladas para dar cuenta de la cuestión de la intencionalidad para estados con una descripción computacional.

Por otro lado, si bien el PNC pretende ser una reconstrucción realista de una tendencia en el ámbito de las semánticas para los contenidos mentales, lo que se intenta respetar son las tesis básicas de las teorías y no las posiciones personales de sus autores con respecto a ciertos temas (esto se verá más claramente en §3).

<sup>2</sup> A pesar de que las propiedades intencionales son una subclase de las mentales, los filósofos suelen utilizar indistintamente los términos “mental” e “intencional”, y también “cognitivo” y similares (ver nota 28 del capítulo 4). Si bien no acuerdo con esta práctica dado que, en parte, es responsable (o quizás, la consecuencia) de las confusiones que intento aclarar en esta tesis, la respeto a lo largo de la misma porque en estos términos se han planteado las cuestiones que nos ocupan. Toda vez que me aparto de esta práctica en función de mostrar lo inconducente de la misma lo dejaré sentado explícitamente.

o bien nos quedamos con un dualismo no deseado. De manera que, si además de ser realista intencional se es fisicalista, esto es, se cree que las propiedades básicas y últimas son las físicas, entonces el problema a resolver para un naturalista es cómo encajar las propiedades semánticas/ intencionales en un mundo cuyas propiedades son físicas. El naturalismo no es más que la consecuencia metodológica del fisicalismo (cf. Fodor 1994, p. 5) que lleva básicamente a naturalizar aquello que en principio no parece ser *per se* parte del mundo natural. Así, las propiedades semánticas/ intencionales o el contenido de los estados mentales o lo intencional, o lo que sea mental, tiene que poder relacionarse con alguna propiedad del mundo físico.<sup>3</sup>

El proyecto de naturalización de las propiedades mentales asimila la idea de explicación en términos naturales a la de reducción (parcial, si se trata sólo de dar condiciones suficientes). La búsqueda de condiciones suficientes naturales para que algo físico sea intencional implica que lo intencional queda reducido a (explicado a partir de) lo físico, pero no implica la eliminación de las nociones intencionales (en el plano semántico) o del fenómeno de la intencionalidad (en el plano ontológico), sino que es considerada como una manera de conservarla/lo. En otras palabras, es una estrategia que permite que lo no-mental explique lo mental de modo de que lo mental no quede fuera del ámbito de lo natural.

El mundo natural al que se alude en este proyecto de naturalización es el de las ciencias naturales, por eso el segundo calificativo de *cientificista*.<sup>4</sup> Si bien tanto el naturalismo científicista como el no científicista consideran que no hay estados o propiedades mentales que 'flotan libremente', el primero cree que la ciencia puede dar cuenta de ellos. Así, sean las que fueren las entidades que se postulen, éstas, de alguna manera, se relacionan con las entidades que las ciencias naturales dicen que hay. Sin embargo, y paradójicamente, sigue en pie el problema heredado del dualismo cartesiano: ¿cómo se conecta lo mental con lo físico? Un naturalista es alguien que acepta que el problema heredado es un problema, pero lo reformula. Se trata, ahora, de la cuestión de cómo algo físico o biológico o natural (sea un individuo, un sistema

---

<sup>3</sup> Parece ser que la tesis metodológica del naturalismo lleva a una restricción ontológica que no es sólo propia del ámbito de lo mental sino que afectaría a cualquier ciencia especial, esto es, a todas las ciencias menos la física. Por ejemplo, un geólogo tendría que explicar que algo físico sea una montaña. Véase Fodor (1994), p. 5.

cognitivo o los estados de un organismo), puede tener propiedades no-físicas, esto es, propiedades que las ciencias naturales no parecen admitir en su ontología.

Lo que ofrece, entonces, un naturalista científicista que cree en el proyecto de naturalizar el contenido de los estados mentales es una teoría formulada en términos no-mentalistas, en la que figuren sólo las propiedades que las ciencias naturales estén dispuestas a admitir. No es sorprendente advertir que lo que en la epistemología constituyó el proyecto de "abandono de la meta de una filosofía primera" (Quine 1975, p. 92) se transforma en esta metafísica científicista de la filosofía de la mente, en el proyecto de elevar a las ciencias naturales al rango de filosofía primera.<sup>5</sup>

A continuación, me ocuparé de desarrollar la semántica informacional (§1) y la telosemántica (§2) de manera de explicitar las tesis básicas del PNC (§3). En el capítulo siguiente discutiré el PNC y propondré una manera de entender el fenómeno de la intencionalidad basado en lo que denomino un naturalismo amplio.

## 1. La semántica informacional

El objetivo de una semántica para los estados intencionales no es el de dar cuenta de por qué una persona tiene las creencias que tiene o cómo se forman las creencias, sino en qué consiste tener un contenido intencional y qué es lo que hace que una creencia tenga un contenido determinado, por ejemplo, que sea acerca de las vacas y no acerca de los perros. Uno de los intentos para resolver la cuestión de cómo algo físico puede tener propiedades no-físicas, y contenidos intencionales específicos, son las teorías covariacionales del contenido, también llamadas semánticas de indicadores o semánticas informacionales.

---

<sup>4</sup> Sigo a Pettit (1992) en la utilización de esta expresión.

<sup>5</sup> El proyecto de abandono de una filosofía primera en el ámbito de la epistemología significaba oponerse al carácter fundamentador de la filosofía. Si bien este proyecto naturalista no dice nada acerca de la epistemología, adopta y extrema una de las consecuencias ontológicas del proyecto



La idea central de las teorías covariacionales es que las representaciones mentales covarían nomológicamente con cosas (eventos, propiedades, etc.) del mundo, dadas ciertas condiciones naturales. Estas teorías proporcionan una especificación reduccionista (si ofrecen condiciones necesarias y suficientes) o parcialmente reduccionista (si sólo ofrecen condiciones suficientes) de hechos semánticos/intencionales (el contenido de los estados mentales) en términos de hechos físicos (condiciones físicas bajo las cuales un estado físico tiene contenido). Las distintas propuestas covariacionales suelen diferir en la especificación de las condiciones naturales para que un estado mental tenga un contenido específico.

La semántica informacional de Fodor incorpora la idea básica de la Teoría Informacional de F. Dretske (1981) y le adiciona su propuesta que se denomina *Teoría de la Dependencia Asimétrica* (TDA, Fodor 1987, cap. 4; 1990b; 1994). Veamos en primer lugar en qué consiste, de manera resumida, la teoría de Dretske de modo de tener en claro lo que la TDA le agrega. La semántica de Dretske es un intento de aplicar la teoría matemática de la información de Shannon (1948) para dar cuenta del desarrollo de sistemas cognitivos a partir de mecanismos puramente físicos que procesan información. La noción básica es la de información, y dado que es una noción objetiva puesto que no supone ningún proceso de interpretación por parte de un agente cognitivo sino que se define en términos de relaciones nómicas entre eventos (casos o tipos), se la puede utilizar para explicar la noción de representación en términos naturalistas.

Dadas ciertas relaciones nómicas entre una señal (o indicador) y aquello que la causa, es posible inferir que la señal *lleva información acerca de* aquello que la causó. La relación semántica fundamental que da cuenta de la intencionalidad es la de *llevar información acerca de* que se explica en términos naturales de *estar conectado nómicamente con*. En este sentido, la intencionalidad es ubicua, no sólo los estados mentales poseen la propiedad de ser intencionales sino muchas otras cosas, como el lenguaje, los artefactos, los organismos vivos, y demás. Desde esta perspectiva, un sistema intencional es cualquier sistema físico cuyos estados internos dependen de

---

naturalista epistemológico, esto es, la idea de las ciencias naturales como piedra de toque de la ontología.

manera legaliforme de los valores de magnitudes externas (Dretske 1980). Así, llevar la información de que  $p$  es una propiedad natural relacional mediada por leyes. Por ejemplo, si una bombilla de luz se enciende cuando se presiona una perilla, el estado de estar encendida la bombilla lleva la información de que se presionó la perilla. O, las manchas rojas en la cara de un niño llevan la información de que tiene sarampión porque hay una relación nómica entre las manchas rojas y el sarampión.

Dado que la intencionalidad es una relación natural, lo que hay que explicar no es tanto la cuestión general de cómo algo físico puede tener la propiedad de ser intencional (contenido informacional) sino la cuestión particular de cómo algo físico puede tener ciertos contenidos específicos, es decir, cómo un sistema físico puede tener el tipo de mente que tenemos los humanos (contenido mental). Se puede entender la distinción entre el contenido informacional y el mental utilizando la distinción subpersonal-personal que hemos visto en el capítulo anterior (§1). Los agentes cognitivos son sistemas que procesan información, gran parte del procesamiento se realiza en el nivel subpersonal, en el que los mecanismos operan sobre el contenido informacional, sin que el sujeto (ni los subsistemas) tenga acceso conciente a la información (en el sentido de que no la conocen o entienden). En cambio, los contenidos mentales pertenecen a estados de pensamiento (APs) y experiencias que guían la conducta y a los cuales el sistema global tiene acceso conciente, y en este sentido pertenecen al nivel personal.

Así, el contenido informacional y el mental son distintos tanto en intensión como en extensión, pero ambos son intencionales en el sentido de que sus contenidos se fijan en función de aquello que los causa, más específicamente, por una relación causal entre algo en el mundo y una representación (que lleva información acerca de eso en el mundo). La propuesta de Dretske consiste entonces en construir un concepto de contenido mental a partir del de contenido informacional que sea coextensivo con el concepto tradicional de contenido mental (esto es, que sus instancias coincidan en las condiciones de verdad o que denote una propiedad nómicamente equivalente a la denotada por el término tradicional).

La construcción de la noción de contenido mental se realiza en tres etapas (Cf. Woodfield 1986). En primer lugar, Dretske distingue entre información analógica y

digital para dar cuenta de que el contenido mental a diferencia del informacional es unívoco, así de la múltiple información que puede llevar el contenido informacional es posible abstraer la información unívoca del contenido mental. En segundo lugar, distingue entre estructuras cognitivas y estructuras no-cognitivas para dar cuenta de que no todo estado informacional da lugar a un estado mental. Y, por último, ofrece una teoría de la representación errónea para dar cuenta de la cuestión de que el contenido mental, a diferencia del informacional, puede ser falso, esto es, puede representar el mundo de una manera en la que no es.

Estos tres pasos han recibido críticas pero fundamentalmente es la teoría de la representación errónea la que ha sido el foco de las objeciones. Cualquier teoría naturalista que no resuelva satisfactoriamente este problema no puede contar como una respuesta a la cuestión de cómo los estados mentales poseen el contenido específico que tienen. Dretske (1981) sostiene que un estado mental lleva la información de que  $p$  dadas ciertas condiciones óptimas y éstas aluden a las condiciones durante el período de aprendizaje. Así, se identifica las condiciones de verdad de una creencia con las que la causaron durante el período de aprendizaje (y de formación de creencias), de manera que las causas que podrían dar lugar a esa creencia luego de ese período no formarían parte de sus condiciones de verdad y por ende, llevarían a creencias falsas.

Fodor (1990a, p. 61 y ss.) objeta a esta propuesta de Dretske que no es posible establecer un punto en que se detiene la formación de creencias, que se aprende un concepto igualmente robusto aunque se esté expuesto a condiciones erróneas (más adelante vuelvo sobre la cuestión de la robustez), y que esta propuesta no se puede aplicar a creencias innatas. Dretske (1986) intenta un enfoque diferente introduciendo aspectos teleológicos. Lo que representa un estado interno está determinado por la función para la cual fue diseñado evolutivamente el mecanismo cognitivo. Así, algunas estructuras, como los estados mentales, tienen la función de recoger información del entorno para cubrir necesidades biológicas, de manera que la función de sus contenidos es indicar (llevar información acerca de algo en el mundo). Un contenido en particular se especifica por lo que indica de hecho cuando los mecanismos cognitivos funcionan normalmente y las

circunstancias son normales. Así, los casos de representación falsa se deben al mal funcionamiento de los mecanismos cognitivos cuando las circunstancias no son las adecuadas.

Esta propuesta teleológica también tiene dificultades, básicamente por las razones que da Fodor: la noción de función lleva a la indeterminación del contenido debido a la indeterminación misma de cualquier función (Fodor 1990a), y la noción de condiciones normales lleva a un círculo puesto que para explicar el contenido se apela a esta noción pero para fijar las circunstancias se requiere saber cuál es el contenido (Fodor 1987, cap. 4; Loewer 1987). Dado los problemas que parece tener la concepción teleológica a la hora de dar cuenta de la determinación del contenido, Fodor (1990b) prefiere una concepción no-teleológica de la semántica informacional.<sup>6</sup>

Fodor adopta de Dretske (1981) la idea fundamental de la semántica informacional de que el contenido está determinado por relaciones nómicas y que el contenido es información, pero no asimila completamente contenido a información, esto es, no asimila completamente el contenido de la representación a la información que lleva. La relación semántica fundamental es la de *expresar una propiedad* (a diferencia de la dretskeana de *llevar información acerca de*) que se explica en términos de relaciones nómicas entre instancias en el mundo de la propiedad expresada por una representación e instancias en el organismo de esa representación.

Siguiendo la idea fundamental de la semántica informacional, la covariación es una relación nomológica/causal entre propiedades, esto es, entre la instanciación de una cierta propiedad (en el mundo) y la instanciación de una representación (en el organismo) que es acerca de esa propiedad.<sup>7</sup> Por ejemplo, cuando se piensa VACA,

---

<sup>6</sup> Fodor (1990d) flirteó con una solución teleológica al problema del error (apelando a condiciones óptimas) que abandonó en (1990b). El relato de la evolución de su postura se puede ver en Fodor (1990a).

<sup>7</sup> Fodor utiliza la noción de causa como la relación natural que explica la relación de representación. En el único lugar donde la explicita sólo dice que para que haya una relación causal entre dos cosas "algo tiene que ocurrir *en el mundo*" (1990c, p. 189). No la analiza en profundidad puesto que considera que no es necesario ofrecer una teoría acerca de la relación causal. En todo caso, la semántica estaría utilizando el mismo aparato metafísico que el resto de las ciencias empíricas, no es parte de su tarea ofrecer una teoría de la causalidad.

este concepto significa *vaca* porque hay una ley que conecta esta instancia de representación con la propiedad de *ser vaca* (la ley 'las vacas causan VACAS'), apoyada por contrafácticos (por ejemplo, 'si se estuviera en presencia de vacas, éstas causarían pensamientos acerca de las vacas').<sup>8</sup>

De esta manera, el contenido surge y se determina por relaciones nómicas puramente contrafácticas que se expresan con condicionales subjuntivos del tipo si las circunstancias fueran (o hubieran sido) tales y cuales, pasarían (o hubieran pasado) tales y cuales cosas. Así, una representación tiene el contenido tal y cual porque si se dieran instancias de tal y cual, éstas causarían instancias de esa representación. Lo que interesa para la determinación del contenido no son las covariaciones causales efectivas entre las propiedades subsumidas bajo las leyes sino las covariaciones que habría dadas ciertas circunstancias.

Hasta aquí las representaciones que constituyen el contenido son causadas por propiedades del entorno y llevan información acerca de lo que las causó. Sin embargo, esta teoría causal cruda tiene un problema: hay muchas cosas en el entorno que pueden causar una representación determinada y puede ocurrir que, en algunos casos, esa representación no sea acerca de aquello que la causó. Utilizando nuestro ejemplo, no todas las cosas en el mundo que causan instancias de VACA constituyen su contenido y a la vez VACA puede ser causado de muchas maneras (puede haber muchas razones por las que uno piense en vacas) y, sin embargo todas significan (o tienen el contenido) *vaca*.

Esto es así porque el ámbito de los contenidos semánticos es mucho más acotado que el de las relaciones causales. La información es sensible a la etiología mientras que el contenido no: si algunas instancias de VACA son causadas por vacas y otras instancias de VACA son causadas por no-vacas, las primeras llevarán información acerca de las vacas, mientras que las segundas no. En cambio, el contenido de VACA (*vaca*) es algo que se supone que todas las instancias de VACA comparten con independencia de la heterogeneidad de sus causas (esto es lo que se

---

<sup>8</sup> Los nombres de los conceptos están escritos en mayúsculas y los nombres de sus valores semánticos en itálicas (por ejemplo, 'VACA' nombra el concepto que expresa la propiedad de *ser vaca*). Los nombres de las expresiones en español aparecen entre comillas.

llama la 'robustez' del significado). El significado es robusto mientras que la información no, la información es ubicua dado que estamos rodeados de relaciones covariacionales fiables, mientras que no hay significado en todas partes.

La TDA es introducida como un intento de solucionar el problema tradicional de las teorías causales de que no todo lo que causa (o si se prefiere, lo que entra en relaciones nómicas con) una instancia representacional es semánticamente relevante para su contenido. La idea fundamental es que las leyes causales que determinan el contenido de las representaciones mantienen entre sí relaciones de dependencia (empírica) asimétrica: "Si las muestras de "A" que son causadas por Bs ... representan falsamente los Bs como As-, entonces *habría* un sendero causal desde los As a los "As", aunque no hubiera un sendero causal desde los Bs a los "As". Pero no habría un sendero causal desde los Bs a los "As" si no hubiera un sendero causal desde los As a los "As" (Fodor 1987, p. 160). Esto quiere decir que las conexiones B-"A" encontradas en el mundo efectivo dependen asimétricamente de las conexiones A-"A" encontradas en el mundo efectivo, de manera que si se quiebran las conexiones A-"A", se quebrarían las conexiones B-"A", pero no a la inversa.<sup>9</sup>

Siguiendo el ejemplo de Fodor, VACA tiene el contenido *vaca* y no *caballo en una noche oscura* porque el que haya instancias de VACA causadas por caballos en una noche oscura depende de que haya instancias de VACA causadas por vacas, de manera que la ley de que los caballos en noches oscuras causan VACAS depende de la ley de que las vacas causan VACAS pero no a la inversa, porque si no hubiera una conexión causal entre VACA y vaca no habría una conexión causal entre VACA y caballo en una noche oscura.<sup>10</sup> La dependencia entre las leyes es sincrónica: caballo en una noche oscura no causaría VACA *ahora* a menos que vaca causara VACA *ahora*, esto quiere decir que no se requiere de la existencia o instanciación

<sup>9</sup> Cuando se habla de "romper la conexión A-"A"" se alude específicamente a la conexión entre la propiedad en virtud de la cual los As causan "As" (la propiedad de ser A) y la propiedad de ser la causa de los "As" (Fodor 1990b).

<sup>10</sup> Fodor (1987, p. 157) recoge así la vieja idea platónica de que las falsedades dependen ontológicamente de las verdades.

previa de la conexión vaca-VACA a la conexión caballo en una noche oscura-VACA.<sup>11</sup>

El contenido se constituye, entonces, por la información (covariación causal fiable, esto es, apoyada por contrafácticos) más la dependencia asimétrica (o robustez, que permite abstraer el contenido de la información). De manera que, las condiciones suficientes naturales para que surja el contenido son un conjunto de relaciones nómicas entre propiedades que mantienen entre sí una dependencia asimétrica y que están expresadas en términos no intencionales:

[El símbolo] "X" tiene el significado X [una propiedad] si:

1- *Condición de información*: 'Los Xs causan "Xs"' es una ley.

2- *Condición de dependencia asimétrica*: Para todo Y no igual a X, el que los Ys *qua* Ys causen efectivamente "Xs" es asimétricamente dependiente de que los Xs causen "Xs". (Fodor 1990b, p. 121)<sup>12</sup>

La semántica naturalizada consistiría en "una teoría que articule, en términos que no sean ni semánticos ni intencionales, condiciones suficientes para que un fragmento del mundo *sea acerca de* (exprese, represente, sea verdad de) otro fragmento" (Fodor 1987, p. 146). El objetivo de la TDA es así explicar la relación 'R representa S' es verdadero si C, donde C alude a condiciones naturales, no expresadas en términos semánticos o intencionales. Las condiciones 1 y 2 se expresan en términos naturales de relaciones causales y expresiones que denotan

<sup>11</sup> La dependencia asimétrica es sincrónica porque depende de las disposiciones efectivas del organismo. En el caso ya clásico de cuál es el contenido de la representación de una rana cuando saca su lengua para atrapar comida (ya que parece ser la disyunción de 'mosca o punto negro en movimiento'), la TDA sostiene que el contenido es *punto negro en movimiento* dado que la ley 'las ranas atrapan moscas' depende asimétricamente de la ley 'las ranas atrapan puntos negros en movimiento'. La cuestión no es diacrónica, no es una cuestión de que las ranas no hubieran desarrollado la disposición de atrapar puntos negros en mundos donde los puntos negros no fueran moscas. Es una cuestión de que las ranas intentarían atrapar puntos negros en mundos donde hay puntos negros pero no moscas, y no a la inversa. Cf. Fodor (1990b) y (1987), pp. 160-1.

<sup>12</sup> En la formulación de Fodor (1990b, p. 121) figura además la condición histórica-causal: algunos "Xs" son efectivamente causados por Xs (la historia efectiva de "X" es constitutiva de su significado). No la incluyo como parte de las condiciones suficientes para el contenido porque en Fodor (1994) se vuelve a la formulación originaria puramente contrafáctica de Fodor (1987).

objetos intensionales (como leyes, contrafácticos y propiedades, Fodor 1990b, p. 127). El proyecto de naturalización (cualquiera sea la teoría sustancial) no excluye la utilización de nociones intensionales como las teleológicas o los contrafácticos (Cf. Loewer 1987).<sup>13</sup> Fodor prefiere nociones como esta última a nociones teleológicas por las razones que hemos visto y las que veremos en el capítulo siguiente.

La condición 1 supone establecer las circunstancias que si se obtuvieran permitirían que una instanciación de una propiedad (por ejemplo, la de *ser vaca*) cause la instanciación de una representación mental (por ejemplo, VACA), es decir, supone establecer los contrafácticos (las vacas causarían/habrían causado VACAS si...). Esto no es una tarea fácil pero habría circunstancias psicofísicas, que relacionarían al organismo con el entorno, especificables en términos no-semánticos ni intencionales ni teleológicos. Fodor (1987, p. 166; 1990c, p. 48, n. 1) da algunos ejemplos como el vocabulario de las longitudes de onda, la potencia de la luz, las irradiaciones de la retina y cosas por el estilo y afirma que, en última instancia, es una cuestión de intuición puesto que tampoco hay una manera satisfactoria de decir cuáles términos son semánticos o intencionales.

Hasta aquí se han desarrollado aquellos aspectos de la semántica informacional de Fodor que servirán para establecer las tesis básicas del PNC. Por supuesto, hay algunas cuestiones que hubieran requerido de un desarrollo mayor y hay otros temas relacionados que no he podido abordar. Sin embargo, esto no será un obstáculo para la comprensión del PNC dado que el objetivo de esta reconstrucción ha sido la familiarización con los aspectos básicos que conformarán las tesis de este proyecto. A continuación se hará lo propio con la teleosemántica.

## 2. La teleosemántica

---

<sup>13</sup> El naturalismo (fiscalista), a diferencia del materialismo, no niega independencia ontológica a las entidades abstractas. Cf. Fodor (1990a), p. 132, n. 6.



Tanto la semántica informacional como la teleosemántica se enmarcan en la concepción funcionalista de los estados mentales. Tal como hemos visto en el capítulo 1 (§ 2.2), los estados mentales son estados que se caracterizan por el rol funcional que poseen. Las típicas teorías funcionalistas, que son las que se suelen desarrollar en los textos filosóficos que presentan la concepción funcionalista, utilizan un concepto disposicional-causal de función: las funciones de los estados mentales se conciben como disposiciones con base causal, de manera que los estados mentales se especifican en virtud de los roles causales efectivos que poseen, esto es, lo que hacen de hecho (o tienen capacidad, de hecho, para hacer). Siguiendo las presentaciones habituales del funcionalismo, éste fue el concepto de función que se adoptó en general en el capítulo 1. Sin embargo, la noción de función puede ser entendida en términos teleológicos, esto es, apelando al concepto 'biológico' de función. Así, las funciones de los estados mentales se conciben desde la perspectiva de sus propósitos o aquello que se supone que deben hacer.

Aplicado esto último a los contenidos de los estados mentales, la idea fundamental de la teleosemántica es que se puede explicar la intencionalidad en términos naturalistas por medio de un concepto teleológico de función.<sup>14</sup> La que se ha ocupado de desarrollar, en detalle, esta perspectiva desde su libro *Language, Thought, and Other Biological Categories* es la filósofa Ruth Millikan. Otros filósofos han propuesto teorías dentro de esta línea teleológica como David

---

<sup>14</sup> El panorama de las relaciones entre las concepciones acerca de los estados mentales y el contenido es más complicado de lo que las clasificaciones de teorías hacen suponer. Los defensores de las teorías informacionales, suelen sostener una concepción funcionalista no-teleológica de los estados mentales. Sin embargo, las teorías informacionales, que son explicaciones no-teleológicas del contenido, pueden adosar elementos teleológicos (en particular para dar cuenta del problema del error), como es el caso cuando apelan a circunstancias normales (Dretske 1986) u óptimas (Fodor 1990d), así la explicación tiene la siguiente forma básica: si es una ley que tal y cual cosa en el mundo causa una instancia de una representación en circunstancias normales u óptimas, entonces, la representación tiene el contenido tal y cual. Cuando Fodor (1990b) agrega la TDA a la semántica informacional para dar cuenta de manera completa del contenido abandona cualquier aspecto teleológico.

Millikan (2001) considera que el enfoque teleológico puede adosarse a cualquier semántica básica de las representaciones mentales, como la semántica informacional, la semántica de rol conceptual o la teoría del isomorfismo. En el caso de las dos semánticas del contenido que nos ocupan, la teleosemántica, como veremos más adelante, se ancla en la teoría del isomorfismo mientras que la TDA, como hemos visto, se ancla en la semántica informacional (en este sentido sería más apropiado denominar TDA a la teoría de Fodor para diferenciarla de la teoría base que es

Papineau (1984, 1987, 1990), Colin McGinn (1989) y Andrew Woodfield (1990). La teoría de Millikan contiene un aparato conceptual muy rico en terminología técnica, y por ende constituye una teoría sumamente compleja, de manera que intentaré desarrollarla acudiendo sólo a aquellos conceptos que sean imprescindibles para lograr una comprensión adecuada de esta propuesta.

La teleosemántica se construye a partir de la noción técnica de 'función propia' que fue acuñada por Millikan (1984) para dar cuenta de las analogías (y diferencias) que hay entre distintos ítems pertenecientes a distintas categorías como cuerpos orgánicos, elementos del lenguaje, herramientas, conductas con propósito, representaciones internas, costumbres, y demás. La idea es que un ítem *A* tiene la función propia *F* si se da alguna de las siguientes condiciones (necesarias y casi suficientes) : 1. *A* se originó como una reproducción (o copia) "de algún/os ítem/s anterior/es que, *debido* en parte a la posesión de las propiedades reproducidas, han realizado de hecho *F* en el pasado, y *A* existe por (un 'por' histórico-causal) esa/s realización/es", 2. *F* se deriva de una *F* originada tal como se formula en 1 (Millikan 1989b, p. 288).

Así, todo ítem tiene una función propia que consiste en los efectos que han permitido la selección de sus ancestros para la reproducción (o copia) y con ello la supervivencia. La función propia de un elemento depende de su historia, del hecho de que tuvo que tener ancestros que hayan ejecutado esa función propia, y no de los poderes causales o la constitución física de ese elemento. Los ancestros determinan si algo tiene una función determinada dado que su ejecución en el pasado llevó a la conservación de ese elemento, independientemente de si ese elemento realiza la función actualmente. Puede ocurrir que un organismo 'defectuoso' no pueda realizar su función propia, esto no quiere decir que no la posea, por ejemplo, la función propia del corazón es bombear sangre al cuerpo (porque este es el rasgo que permitió que los corazones se reprodujeran) aunque un corazón enfermo no lo pueda hacer. Las funciones propias son una cuestión de los efectos históricos de instancias pasadas de un cierto tipo que permiten dar cuenta de la existencia de instancias presentes.

---

la semántica informacional. Sin embargo, no haré esta distinción dado que el mismo Fodor suele llamar a su teoría simplemente semántica informacional).

La noción de 'propia' recoge el sentido latino de 'propiedad', así, la función propia es la que se posee (por pertenecer a una familia reproductiva/copiadora de elementos similares). Si bien es una noción normativa dado que establece un estándar a partir del cual medir el alejamiento del mismo (y en este sentido, se puede establecer cuando un mecanismo no funciona apropiadamente), no es evaluativa ni prescriptiva. El sentido normativo se relaciona con la idea de que la realización de una función depende, en parte, de lo que *se supone* que un sistema hace o *debe* hacer o *está diseñado para* hacer (por eso, la función propia de un elemento alude al propósito natural de ese elemento, cf. Millikan 2000).

Como veníamos diciendo, hay varias cosas que tienen función propia, no sólo aquellas que se reproducen genéticamente (por los mecanismos de la selección natural darwiniana) como los genes, los estados mentales, y las conductas no aprendidas. Hay muchas cosas que se reproducen en el sentido de que se copian en virtud de los efectos que tienen. Por ejemplo, los artefactos son copias de ejemplares anteriores que se han perpetuado en virtud de sus efectos. También las conductas aprendidas por ensayo y error o por imitación o incluso las convencionales (costumbres, formas del lenguaje, moda, y similares). Todos estos ítems tienen ancestros que se reprodujeron (se copiaron) en virtud de ciertos efectos propios. En este sentido, la noción de función propia no fue pensada primordialmente para aplicaciones estrictamente biológicas. Sin embargo, las categorías no-biológicas pueden comprenderse en analogía con la noción 'biológica' de función propia y en este respecto son categorías biológicas en un sentido amplio o metafórico (Millikan 1999a; 2002).

Si bien no se hace un uso literal del término 'biología' para el caso de algunos elementos (por ejemplo, los artefactos), se desprende de la teleosemántica que el paradigma para entender la intencionalidad de los estados mentales es la biología. En este sentido, Millikan dice que "la clasificación psicológica es clasificación biológica ... categorías tales como creencia, deseo, memoria, percepto, y conducta intencionada son categorías de función biológica ... las categorías o tipos como la creencia de que está lloviendo, el deseo de visitar París, el percepto de un gato, y el disparo

intencionado de un conejo se esculpen con referencia a funciones biológicas” (1993b, pp. 212-13).

En tanto que los estados mentales responden a una categoría biológica en sentido estricto, tienen una función biológica propia. Por ejemplo, las creencias tienen la función biológica propia de contribuir a los procesos de inferencia práctica que sirven para cumplir los deseos y éstos tienen la función biológica propia de producir una correspondencia entre las representaciones y el mundo (Millikan 1986; 1989a). Los procesos de selección natural que reprodujeron estos ítems biológicos no consisten, en este caso, en la transmisión genética de las creencias, deseos, y similares sino en la adquisición de conceptos por medio de mecanismos de aprendizaje. Así,

Si la representación incluye una actitud proposicional particular de un humano en particular, la derivación [de la explicación del contenido de esa representación] llegará retrospectivamente, al menos, hasta la historia evolutiva de esos mecanismos cognitivos humanos que son responsables de la formación de conceptos, esto es, de las capacidades humanas de aprender cómo aprender. (Millikan 1995, p. 287)

Ahora bien, la realización de una función depende, además del elemento normativo del diseño (de lo que se supone que un sistema hace), de las condiciones Normales (adecuadas u óptimas) para un funcionamiento exitoso de los mecanismos. Las condiciones Normales son las históricamente óptimas bajo las cuales una función se realizó. Esto quiere decir que una condición es la adecuada para la realización de una función no por ser la condición actual (o posible) bajo la cual un sistema ejecuta su función sino por ser la que permitió que ese sistema reprodujera una capacidad con esa función. De manera que no se trata de una cuestión contrafáctica, en la que dadas las condiciones adecuadas el sistema realizaría cierta función (como en la semántica informacional) sino que el sistema debe realizar esa función, para ello fue diseñado (Cf. Gomila 1995).

La mayúscula en 'Normal' alude al carácter no efectivo ni frecuencial ni promedio de las condiciones sino a su carácter normativo (para lo que fue diseñado) e histórico (relativo a la historia de una función). Algunas funciones rara vez se realizan porque las condiciones son poco frecuentes o se realizan bajo condiciones raras. Un ejemplo de lo primero es la función de los espermatozoides de fecundar el óvulo, muy pocos la realizan. Un ejemplo de lo último es el reflejo vomitivo, una condición Normal para vomitar es la presencia de veneno en el estómago, sólo bajo esa condición es que la función de evitar que el cuerpo se intoxique tuvo históricamente efectos beneficiosos. Las condiciones Normales son las que hay que mencionar para dar una 'explicación Normal' del buen funcionamiento de un sistema. La explicación Normal es la que prepondera en los casos históricos en los que se realizó esa función y da cuenta de su ejecución diciendo cómo funcionó adecuadamente a lo largo de su historia, mencionando leyes naturales y la estructura del sistema.

A partir de la noción de función propia se define la noción de 'ícono intencional' que es la utilizada para dar cuenta del fenómeno de la intencionalidad (Millikan 1984, cap. 6; 1993a). Las condiciones para que algo sea un ícono intencional son las siguientes: [1] tener una función propia (por ser miembro de una familia reproductiva), [2] estar situado entre dos mecanismos cooperativos, un mecanismo que los produce y otro que los consume, coordinados entre sí de manera que cada uno es la condición Normal para que el otro ejecute su función, [3] servir para adaptar el mecanismo consumidor o intérprete a condiciones del entorno de manera que éste pueda ejecutar su función propia, y [4] ser de alguno de estos dos tipos: indicativos e imperativos. Los indicativos aluden a las condiciones Normales bajo las cuales se realiza una función propia y los imperativos se relacionan con la función propia que poseen. Veamos en que consiste cada condición.

[1] Los íconos intencionales ('ícono' en el sentido de C. S. Peirce, cf. Millikan 1984, p. 95) son estructuras o dispositivos que *se supone* que se proyectan en (o se corresponden con, o en sentido más estricto, son isomórficos con) rasgos del mundo a los cuales están adaptados, de manera de ejecutar su función propia. No están preformados en los organismos sino que se adquieren en respuesta al entorno. El

ejemplo paradigmático de un ícono intencional, al que Millikan apela constantemente, es la danza de las abejas cuya función es estimular el sistema nervioso de las abejas que la están mirando para que vuelen en la dirección del néctar. La función propia de la danza podrá realizarse sólo si el entorno se corresponde con el ícono, esto es, si la ubicación del néctar se corresponde con la forma de la danza.

[2] Hay mecanismos en los organismos que producen los íconos intencionales. Por ejemplo, las estructuras cristalinas de los ojos producen las imágenes en la retina, ellas fueron diseñadas evolutivamente para establecer correspondencias entre imágenes retinales y rasgos del entorno. Estos patrones de correspondencia son íconos intencionales producidos por los cristalinos. Lo mismo ocurre en las abejas, tienen mecanismos que producen la danza. Por otro lado, tiene que haber también mecanismos que interpretan los íconos de manera de que éstos sirvan de guía para el comportamiento, esto es, para que el ítem en cuestión realice su función propia. Por ejemplo, las abejas tienen mecanismos que interpretan la estructura danza-localización del néctar de manera de cumplir con la función de volar en dirección del néctar. Los mecanismos productores y consumidores pueden residir o no en el mismo organismo.

[3] Siguiendo con el ejemplo de la danza, la condición a la que se adaptan los mecanismos intérpretes de las abejas que miran la danza es la localización del néctar. Es en virtud de que la localización del néctar es la condición Normal para la ejecución de la función de volar en dirección al néctar que esta función se realiza.

[4] Los íconos son imperativos en virtud de lo que hacen, en otras palabras, dicen qué hacer al mecanismo que los interpreta, mientras que son indicativos en virtud de la condición Normal que permite que hagan lo que hacen, en otras palabras, están determinados por cómo son los hechos. La mayoría de los íconos intencionales primitivos son a la vez indicativos e imperativos, como en el caso de la danza de las abejas.<sup>15</sup> La representación imperativa es la función de la danza que consiste en

---

<sup>15</sup> Los íconos indicativos tienen el estilo de las creencias (expresan un estado de cosas), mientras que los imperativos tienen el estilo de los deseos (expresan un objetivo). Así, la danza expresa la creencia de que hay néctar en cierta dirección o el deseo de que las abejas vuelen hacia esa dirección o ambos. El ícono intencional de la danza es a la vez indicativo e imperativo porque conecta estados de cosas directamente con acciones específicas de las abejas. En cambio, en el caso de los humanos,

estimular el sistema nervioso de las abejas que la ven para que vuelen en cierta dirección y recolecten el néctar, la representación indicativa de la danza es la localización del néctar.

Las condiciones 1-4 dicen cuándo algo es un ícono intencional, sin embargo no dicen acerca de qué son los íconos. Un ícono es intencional, en parte, por lo que hace, por su función propia, pero la función propia del ítem no determina cuál es su contenido. El contenido de los íconos intencionales es la condición Normal para la realización de su función propia. En el caso de la danza de las abejas, la *función* de la danza de las abejas es estimular el sistema nervioso de las abejas que la ven para que vuelen en cierta dirección y recolecten el néctar, el *contenido* de la danza es la localización del néctar.

Así como hemos visto para el caso de la semántica informacional que el contenido se especifica en virtud de relaciones nómicas entre la representación y lo representado de manera que lo representado determina el contenido, en el caso de la teleosemántica el contenido se especifica en virtud de principios de proyección (o correspondencia) entre la representación y lo representado de manera que lo representado determina el contenido. En esta comparación hay diferencias importantes que hacen a la relación de representación y otras no tan obvias que se relacionan con el contenido. Veamos en primer lugar la relación representacional.

La teleosemántica adopta la explicación de la relación representacional que ofrece la teoría del isomorfismo. La noción de representación es la utilizada en matemática: “se dice que una estructura que consiste en un conjunto de entidades abstractas junto con relaciones designadas entre ellas representa otra estructura así, si puede ser proyectada sobre ella uno a uno” (Millikan 2000, p. 5).<sup>16</sup> Los principios de proyección son reglas de correspondencia biunívocas, en el sentido de funciones matemáticas de proyección, entre la representación y lo representado. Cualquier transformación en la estructura representacional varía isomórficamente con la que se da en el estado de cosas en el mundo. Por ejemplo, las variaciones en el ritmo y el ángulo del eje de la danza de las abejas se corresponden con la

---

las acciones no están directamente ligadas con los estados de cosas, de manera que los íconos son o bien indicativos o bien imperativos. Ver nota 20.

distancia y dirección del néctar. En este sentido, tanto la danza como la condición en el mundo representada por la danza están articuladas, la representación completa es la danza-en-un-tiempo-y-un-lugar-con-un-ritmo-y-una-orientación (Millikan 1989a).<sup>17</sup>

El patrón danza-ubicación del néctar está articulado porque está compuesto de elementos invariantes y otros que varían (Millikan 1984, cap. 6; 1989a). Los elementos invariantes son los que todas las danzas tienen en común, por lo cual pertenecen al mismo tipo. Todas producen lo mismo, esto es, una dirección de vuelo en las abejas que las observan, que es una función (matemática) del patrón de la danza. Es por esto que la danza ha sido seleccionada y en este sentido, todas tienen la misma función propia 'relacional' (porque tienen como función causar una dirección de vuelo en relación con la forma de la danza).<sup>18</sup> Los elementos que varían en el patrón danza-ubicación del néctar son los indexicales relacionados con el tiempo y lugar (Millikan 1995).

La danza es representacional porque hay una explicación Normal, de la relación unívoca, que apela a los elementos invariantes del patrón danza-ubicación del néctar a través de la historia evolutiva de los ancestros de las abejas. Así, la relación representacional no requiere de correlaciones nómicas pero tampoco consiste en correlaciones estadísticas azarosas. Las funciones propias suelen fallar porque dependen de condiciones externas para realizarse (y en este sentido, no puede hablarse de leyes, volveré sobre este punto en el §3), sin embargo su realización no es azarosa sino que responde a razones evolutivas (Millikan 2001, 1989a, 1993a; Woodfield 1990; Quesada 1995).

---

<sup>16</sup> La paginación es del artículo on-line en la página web oficial de R. Millikan.

<sup>17</sup> Cuando hay una correspondencia entre dos estructuras, para cada objeto de una hay un solo objeto correspondiente en la otra, lo mismo ocurre con las relaciones y los estados de cosas que son una relación entre una  $n$ -tupla de objetos. Los objetos, relaciones y estados de cosas de una estructura representan estos elementos respectivamente en la otra porque hay isomorfismo de estructuras (no al revés). Los mapas, por ejemplo, son isomórficos con lo que representan y es en este sentido que Millikan (1986, 1990, 1991, 1993a) sostiene que las representaciones son como mapas.

<sup>18</sup> Millikan (1984, p. 107) dice que no sabe cuáles son los elementos invariantes de la danza, pero los elementos invariantes con los que se corresponde son el sol, la colmena, y el néctar. Es decir que las configuraciones de la danza no responden a otros elementos del mundo que no sean esos.



Con respecto al contenido, tanto en la TDA como en la teleosemántica, el contenido está dado por las condiciones de verdad y se establece por relaciones causales entre la representación y lo representado. Pero mientras que en la TDA son las relaciones asimétricas entre los contrafácticos las que lo especifican, en la teleosemántica son las relaciones causales que asociaron la representación con lo representado las que lo especifican, esto es, las que históricamente llevaron a la selección de los ancestros, o en otras palabras, las condiciones Normales para la ejecución de una función propia.

La condición Normal se deriva de las relaciones invariantes entre la representación y lo representado, y de la explicación Normal de esta relación. La explicación Normal del contenido de la danza dice que hay néctar en cierta localización que mantiene relaciones de correspondencia con la configuración de la danza. Si la localización del néctar es errónea (si la danza es una representación falsa), la danza no ejecuta la función propia de guiar a las abejas que la miran hacia la localización del néctar. O, si aún así, las abejas ejecutan su función propia de volar hacia el néctar, es decir si las abejas encuentran miel en una localización que no mantiene estas relaciones de correspondencia con la configuración de la danza, entonces, no realizan 'apropiadamente' su función, esto es, de acuerdo con una explicación Normal, sino que por razones accidentales no relacionadas con su historia evolutiva que llevó a la proliferación de las abejas.

El contenido en la teleosemántica depende, así, de lo representado y de la relación de correspondencia. Su explicación completa sería la siguiente:

El "contenido" de un ícono intencional se describe diciendo qué tipo de estructura o rasgo tendría que haber habido en el entorno del organismo, para que el ícono se corresponda por su regla de proyección, de modo que sus consumidores lo usen exitosamente de la manera normal, esto es, la manera en que históricamente dió cuenta del diseño entrelazado entre el productor, el ícono y el consumidor. (Millikan 1993a, p. 100)

Algo es una representación sólo para un sistema que la interpreta, de manera que son los mecanismos que consumen las representaciones los que determinan cuáles correspondencias son las adecuadas, es decir, cuál es el contenido de la representación. La función de los mecanismos *productores* de representaciones es que éstas se correspondan con el mundo, que la representación (ícono intencional) coincida con lo representado (la condición Normal en el mundo), pero son los *consumidores* los que determinan el contenido, es decir, cuál regla de correspondencia es la adecuada (cuál sea ésta depende de la historia evolutiva del sistema). La semántica que hay que construir es una semántica para el lenguaje de los consumidores que son los que usan el sistema de representaciones (Millikan 2002). Así, la intencionalidad es específicamente una relación entre los consumidores, las representaciones y el entorno (Millikan 1990).

Hemos visto que el conjunto de los ítems que tienen funciones biológicas es mayor que el de los íconos intencionales, para que un ítem sea un ícono intencional tiene que satisfacer las cuatro condiciones mencionadas más arriba (un ejemplo sería el corazón que tiene la función propia de bombear sangre al cuerpo pero no es un ícono intencional). A su vez, si bien hay muchas cosas que satisfacen las cuatro condiciones para ser un ícono intencional, y en ese sentido, la intencionalidad es un fenómeno ubicuo (todo a lo que llamaríamos signo cumple estas condiciones), el conjunto de los íconos intencionales es mayor que el de los íconos intencionales que son representaciones (un ejemplo es la danza de las abejas por lo que veremos a continuación).

Las representaciones son aquellos íconos intencionales que tienen como una de sus funciones participar en inferencias mediatas, esto es, combinarse con otros íconos para producir íconos con información nueva. Para que esto sea posible, las representaciones tienen que ser identificadas, en el sentido de que el portador de las representaciones tiene que saber cuáles son los referentes de las mismas (Millikan 1984, p. 96). Según las palabras de Millikan:

El acto de identificación ocurre siempre que se juntan dos íconos, ya sean indicativos o imperativos, que se solapan en contenido, para producir o bien otro ícono intencional

o una acción basada en los dos íconos tomados en conjunto. Por ejemplo, el uso conjunto de la vista y la sensación del tacto del mismo objeto para producir una coordinación manipulada del mismo, dando por sentado que la vista y la sensación del tacto dan lugar a íconos intencionales, involucra un acto de identificación, por consiguiente involucra “representaciones” visuales y táctiles. (1993a, p. 121, n. 10)

La danza de las abejas si bien es un ícono intencional no califica como representación porque no es el caso que la danza cumpla con su función apropiadamente cuando un mecanismo intérprete (en las abejas que la miran) identifica los referentes con los que se corresponde. Las abejas realizan su función de volar hacia el néctar sin *saber* acerca de qué son los contenidos de la danza. En cambio, los estados mentales como las creencias están constituidos por representaciones porque la función propia de las creencias (que son íconos indicativos) es participar en procesos de inferencia práctica para el cumplimiento de los deseos (que son íconos imperativos),<sup>19</sup> y para cumplir con su función se requieren actos de identificación (y reidentificación) de sus referentes (Millikan 1984, cap. 8; 1989a).<sup>20</sup>

La función propia del sistema de producción de creencias es producir creencias de que  $p$  si y sólo si  $p$  (esto es, producir una correspondencia entre las representaciones y lo representado), mientras que es el sistema consumidor de representaciones el que establece la relación unívoca que especifica la condición Normal bajo la cual las creencias cumplen con su función propia de contribuir a los procesos de inferencias prácticas. El contenido de las creencias dependerá de las condiciones Normales que permiten cumplir con sus funciones. Las creencias verdaderas son las que se proyectan en el mundo de acuerdo con reglas

---

<sup>19</sup> Cada creencia tiene un rango diferente de funciones propias según cómo se combine en virtud de su forma con otras creencias (Millikan 1989b).

<sup>20</sup> En Millikan (1989a, § v) se enumeran características propias del sistema representacional de los humanos que lo diferencia del de las especies inferiores: 1. no todas sus representaciones son autoreferenciales, 2. las representaciones se almacenan, 3. hay representaciones indicativas e imperativas claramente distinguibles, 4. las representaciones participan en inferencias, 5. sólo los humanos realizan actos de identificación de manera de saber acerca de qué son sus representaciones, y 6. el sistema representacional contiene negación lo cual quiere decir que tiene una estructura proposicional (sujeto-predicado), y en ese sentido, las representaciones tienen contenido proposicional.

especificables (esto sucede cuando las condiciones son las Normales), así un tipo de creencia se corresponderá con algo en el mundo si y sólo si algún mecanismo lo seleccionó porque se correspondió con eso mismo en el pasado.

Las creencias falsas ocurren cuando las condiciones no son las Normales, esto es, las que fueron históricamente cuando tuvieron éxito en ejecutar su función. Al igual que se explicitó para el caso de la danza de las abejas, si una creencia es falsa no cumplirá con algunas de sus funciones propias, si bien puede participar de inferencias no dará lugar a nuevas creencias verdaderas o no ayudará a cumplir deseos. Y, si se diera el caso de que realizara alguna de estas funciones, lo haría por accidente, no de acuerdo con una explicación Normal (Millikan 1986).

### 3. El naturalismo científicista

En lo que sigue intentaré explicitar las tesis principales del PNC en función de las tesis que comparten la TDA y la teleosemántica (y también de aquellas que las hacen ser teorías en competencia por la mejor explicación naturalista del contenido de los estados mentales).<sup>21</sup>

El naturalismo acerca del contenido mental consiste en la siguiente tesis:

(1) *Naturalismo*: las condiciones suficientes para que un estado físico tenga contenido son especificables en términos no-semánticos y no-intencionales.

Tanto Fodor como Millikan ofrecen condiciones suficientes para que un estado físico tenga un contenido intencional. En la TDA, esas condiciones son un conjunto de leyes (causales) que conectan, de manera fiable, instancias de representaciones en el organismo con instancias de propiedades en el mundo, de tal modo que las representaciones llevan información acerca de esas propiedades. A esta semántica

informativa, Fodor le adiciona su TDA para resolver la cuestión de que no todo lo que causa una instancia representacional es semánticamente relevante para su contenido. Así, las condiciones naturales para que surja el contenido son un conjunto de relaciones nómicas entre propiedades que mantienen entre sí una dependencia asimétrica y que son especificables en términos no-semánticos/intencionales.<sup>22</sup>

En la teleosemántica, las condiciones suficientes naturales son un conjunto de condiciones Normales que en virtud de principios de proyección que conectan la representación con lo representado, determinan el contenido. Tanto las relaciones nómicas asimétricas como las correspondencias según las condiciones Normales se formulan en términos no-intencionales, esto es, expresiones en las que no figuran términos como “creencia”, “deseo”, “intención”, y similares, y en términos no-semánticos, esto es, no figuran términos como “significado”, y demás.

Ambas teorías son científicas en el sentido que sostienen la siguiente tesis:

(2) *Naturalismo científico*: los términos no intencionales utilizados en (1) son términos que refieren a propiedades (eventos, procesos, relaciones, y demás) que las ciencias naturales aceptan en sus teorías.

Este es el naturalismo científico propiamente dicho y que da nombre a todo el proyecto dado el peso que tiene esta tesis. Tanto en la teoría de Fodor como en la de Millikan, las condiciones suficientes para que un estado físico sea intencional se formulan en términos aceptados por las ciencias naturales, sólo que ambos apelan a distintas ciencias como los paradigmas que dan cuenta del mundo natural. Tanto Fodor como Millikan son fisicalistas en el sentido de que aceptan que todas las entidades son, o supervienen o se realizan en, entidades físicas. Sin embargo, en cuanto a la explicación de las propiedades intencionales, Fodor apela a la física mientras que Millikan a la biología.

---

<sup>21</sup> Versiones anteriores de las tesis del PNC y algunas de las críticas (en el capítulo 4) fueron desarrolladas en Skidelsky (2002) y en prensa.

<sup>22</sup> En sentido estricto, la condición de información y la de dependencia asimétrica son condiciones necesarias, por separado, y suficientes, en conjunto, para la satisfacción de la condición suficiente para que “X” exprese X. Cf. Fodor 1990b, p. 135, n. 28.

Para Fodor, la relación de intencionalidad, planteada en términos de la conexión nómica entre las representaciones y el mundo, se reduce a una relación física legal, y los términos no-intencionales a los que hay que apelar son del vocabulario de la física (tal como hemos visto en el §1, los ejemplos que da son el vocabulario de las longitudes de onda, la potencia de la luz, las irradiaciones de la retina). En cambio, Millikan sostiene que las condiciones se formulan en términos que no son estrictamente de la física y que aluden a propiedades naturales de la biología. Millikan considera que la psicología es una rama de la fisiología, que a su vez forma parte de la biología y por ende, esta última es la ciencia a cuya ontología hay que apelar. Por eso, la relación de intencionalidad se explica en última instancia en términos evolutivos.

De manera que hay que entender la diferencia en (2) simplemente como la elección de la ciencia paradigmática a la que reducir (en términos de explicar) el fenómeno de la intencionalidad, así las tesis quedan especificadas de la siguiente forma (la "F" es por Fodor y la "M" por Millikan):

(2.F) Los términos utilizados en (1) son términos que refieren a propiedades (u otras entidades) de la *física*.

(2.M) Los términos utilizados en (1) son términos que refieren a propiedades (u otras entidades) de la *biología*.

Las tesis (1) y (2) son tesis filosóficas relacionadas con un ámbito de la explicación mental que es el de los contenidos mentales. No obstante, ambas suponen una tesis naturalista metafilosófica, esto es, acerca de la naturaleza de la empresa filosófica. En la TDA y la teleosemántica hay datos empíricos que intervienen en argumentos que no son ya los tradicionales *a priori*, de modo que, la manera de hacer filosofía no se concibe como una tarea puramente *a priori* sino incorporando los datos, la terminología, los ejemplos y la ontología de las ciencias naturales. Así, por un lado, las tesis filosóficas tienen que ser compatibles con los resultados de las ciencias y, por el otro, la tarea de la filosofía no sólo consiste en la clarificación del

aparato conceptual de las ciencias sino en, por ejemplo, la propuesta de hipótesis empíricas, ambas tareas permitirían integrar, desde un punto de vista empírico y conceptual, el ámbito de lo mental al ámbito de lo físico. Esta idea queda formulada en la siguiente tesis:

(3) *Tesis naturalista metafilosófica*: la filosofía de la mente se desarrolla en conformidad con, o hace uso de, o hace explícito el aparato metafísico de, la ciencia corriente de manera de integrar la mente en el mundo natural científico.

La tesis (3) es implicada por las tesis (1) y (2) dado que no parece plausible que alguien que sostenga (1) y (2) tal como están formuladas, esto es, alguien que lleva a cabo una empresa reduccionista en un sector de la filosofía de la mente, no acepte (3), esto es, que sostenga que la tarea de la filosofía de la mente es exclusivamente conceptual. Por supuesto que la tesis (3) no descarta una tarea conceptual en la filosofía de la mente, pero alguien que sostiene (1) y (2) tal como están formuladas, al menos, se compromete con (3).

La motivación para adherir a las tesis (1) y (2) radica en la conjunción de dos tesis relacionadas con la naturaleza de los estados intencionales. Tanto Fodor como Millikan son realistas intencionales, consideran que hay realmente estados con contenido y que esos estados están conformados por representaciones. Si se parte de una ontología como esta (en la que uno de los términos de la relación intencional son las representaciones y, por supuesto, se descuenta que no se es dualista) entonces, se buscará relacionar las representaciones con algo del mundo natural.<sup>23</sup> Esta idea se puede formular en las siguientes tesis:

(4) *Realismo intencional*: existen estados internos con contenido intencional que es causalmente explicativo de la conducta de un organismo y es semánticamente evaluable.

---

<sup>23</sup> Como este es un punto importante para entender por qué algunos filósofos creen que hay que desarrollar una semántica para el contenido mental y que además sea naturalista, dejaré para el capítulo 4, cuando discuta el PNC, la profundización de la índole de la relación entre el conjunto de las tesis (4) y (5), y el conjunto de las tesis (1) y (2).

(5) *Representacionalismo intencional*: las representaciones mentales son las portadoras de la intencionalidad de los estados intencionales.

Para ambas teorías existen estados intencionales y estos son representacionales. Las representaciones conforman un sistema que el organismo manipula de manera de dar lugar a representaciones nuevas. Este sistema de representaciones guía la conducta de los organismos y ésta se explica apelando a la información que aquellas portan. Ahora bien, para Fodor este sistema representacional tiene propiedades sintácticas (o formales) y semánticas tal como posee el lenguaje natural (aunque no es ninguno de los lenguajes naturales). En la teoría de Fodor, (4) y (5) suponen la hipótesis del lenguaje del pensamiento, que hemos visto en el capítulo anterior (§2.1). En cambio, Millikan (1993a) considera que si bien no sería necesario apelar al formato simbólico de las representaciones dada la teoría del isomorfismo estructural, no descarta la posibilidad de que las representaciones tengan una estructura oracional sólo para el procesamiento cognitivo superior en los humanos. Sin embargo, en este caso las oraciones mentales no responderían estrictamente a la hipótesis del lenguaje del pensamiento dados los varios problemas que esta hipótesis presenta. Veamos sucintamente los argumentos principales que sustentan esta postura.

Tal como hemos visto, la intencionalidad se explica a partir de la relación de correspondencia entre un ícono intencional (elemento que exhibe el rasgo de *ser acerca de*) y un estado de cosas en el mundo. En este sentido, muchas especies tienen intencionalidad sin que sus representaciones tengan estructura oracional (por ejemplo, la danza de las abejas). Por otro lado, hemos visto que un ícono intencional es una representación sólo cuando participa de procesos inferenciales. Los procesos inferenciales a los que se alude son los mediatos<sup>24</sup> en los que se combinan dos ítems que transportan información usando un término medio para producir un tercero que porta información nueva. Los ítems que se combinan no tienen que tener una

---

<sup>24</sup> Los inmediatos son una cuestión tan solo de traducción de información de un medio a otro, por ejemplo de patrones retinales a patrones neurales conexionistas (Millikan 1993a, p. 105).



estructura oracional como cuando se combina información entre perceptos (por ejemplo, un objeto visto y oído en el mismo lugar) para producir una creencia (por ejemplo, que una naranja no está en buenas condiciones).

Las representaciones mentales se definen así en virtud de este procesamiento cognitivo de transformación de la información entre ítems cargados de contenido que no tienen por qué tener estructura oracional. Millikan supone que la mayoría de nuestro procesamiento cognitivo es del estilo ilustrado y en este sentido (y dada la teoría pictórica o de isomorfismo abstracto), el medio representacional habría que entenderlo más como 'mapas', 'modelos' o '*pictures*'. Sin embargo, parte del procesamiento cognitivo superior podría utilizar oraciones mentales para las cuales se requiere, al menos, una estructura sujeto-predicado y que sean pasibles de negación. Aquí sería el lugar natural de la hipótesis del lenguaje del pensamiento.

No obstante, Millikan (1993a) cree que esta hipótesis descansa en una analogía con los sistemas lógicos formales y tiene problemas técnicos que se aplican a todo lenguaje lógico formal. Estos requieren de símbolos que se agrupen en tipos para que se les apliquen reglas formales que determinen la validez de las inferencias, y puesto que los criterios para agrupar las instancias en un tipo particular en los lenguajes formales tienen problemas, no hay manera de que funcione un lenguaje así.<sup>25</sup> La propuesta de Millikan para las oraciones mentales no requiere, por ejemplo, que los elementos del mismo tipo tengan la misma forma para poder realizar transformaciones sobre ellos ni que se les apliquen las mismas reglas de proyección según la forma del elemento porque las instanciaciones de las representaciones se agrupan de acuerdo a su historia evolutiva, a la función propia que comparten (Millikan 1984, pp. 75-76).<sup>26</sup>

---

<sup>25</sup> El criterio no puede ser la 'similitud de las formas' de las instancias puesto que siempre se acude a la intención del que diseña el sistema formal para decir que la instancia *p* corresponde al tipo *P*, y esto es algo que no podría ocurrir en un sistema computacional autónomo. Tampoco el reconocimiento de que ciertas instancias pertenecen a un tipo puede recaer en la 'similitud de reacción' de ciertas partes del sistema. Estos no siempre reaccionan de la misma manera ante el mismo símbolo, sino que la reacción depende del contexto en el que el símbolo esté inmerso (las cadenas en las que participa y sus relaciones con otras cadenas de símbolos). Cf. Millikan (1993a).

<sup>26</sup> Así, la similitud de la forma depende de la similitud de la función propia. Que dos cadenas de símbolos tengan la misma forma (sintáctica) depende de la historia de esos ítems (las instancias de palabras forman parte de familias reproductivas con funciones propias, Cf. Millikan 1984, p. 75). Según las funciones propias de los símbolos, hay formas sistemáticas de producirlos de acuerdo con

Sin embargo, más allá de estos problemas técnicos cabe destacar que la teleosemántica no descarta la hipótesis de un lenguaje del pensamiento (si bien no para todo el procesamiento cognitivo) y que incluso Millikan (2001) considera que esta hipótesis podría ser una de las maneras que podrían adoptar las funciones de correspondencia. Pero lo que importa es que más allá de las discrepancias entre la semántica informacional y la teleosemántica acerca del formato que adoptan las representaciones ambas sostienen las tesis (4) y (5). Ambas sostienen que los vehículos del contenido intencional son objetos internos instanciados físicamente que causan la conducta, y cuyo contenido es proposicional, de manera que puede ser evaluado semánticamente, e intencional.<sup>27</sup>

En ambas teorías, el contenido de las representaciones está determinado por las relaciones (de covariancia legal o de proyección) que mantiene la representación con lo representado. Esto quiere decir que el contenido del estado mental de una persona depende de sus relaciones con el mundo. Esta es la tesis externalista que se opone a la internalista que afirma que el contenido de nuestras creencias no depende de las relaciones de la persona con el mundo. Así, ambos enfoques comparten la tesis del:

---

reglas de correspondencia y hay formas sistemáticas de interpretarlos para que realicen su función propia. De modo que la agrupación en tipos descansa en 'redes de coidentificación' entre los productores y consumidores de las representaciones (Millikan 1993a).

<sup>27</sup> La teleosemántica comparte con la semántica informacional la idea básica de las teorías computacionales de la mente, esto es, los sistemas cognitivos realizan transformaciones de información (computación) entre ítems con contenido (representaciones). Sin embargo, Millikan (1993a) adhiere a una arquitectura computacional conexionista en la que las computaciones no serían sobre símbolos estructurados sintácticamente a la manera de la TRCM, sino sobre representaciones distribuidas que se almacenan de manera superposicional en las redes conexionistas. Es en este sentido que en la nota 1 se hizo notar que según la propia postura de Millikan, la teleosemántica sería apropiada para una teoría representacional/computacional de la mente basada en una arquitectura conexionista, pero que, sin embargo, se suelen considerar a las semánticas de manera independiente de las arquitecturas mentales y en este sentido, han habido intentos de compatibilización de la TRCM con semánticas teleológicas.

Para el caso de una arquitectura conexionista, en particular, las tesis 1.4 de la TRCM acerca de la naturaleza de las representaciones y la tesis 2.2 acerca de los procesos computacionales (§ 2.1 del capítulo 2) adoptarían las siguientes formulaciones: las representaciones son patrones distribuidos de activación sobre nodos de una red que se corresponden con los subsímbolos constituyentes de los símbolos de la TRCM, y los procesos computacionales consisten en algoritmos (numéricos/estadísticos) de propagación de la actividad de los nodos.

(6) *Externalismo*: el contenido de los estados intencionales de una persona está determinado por sus relaciones con el mundo.<sup>28</sup>

La tesis (6) no está asociada a ninguna de las tesis anteriores, es decir, se puede ser externalista sin sostener el naturalismo especificado en la tesis (1)-(3) y sin adherir al realismo intencional/representacional de las tesis (4) y (5). Con respecto a lo primero, hay posturas externalistas que no adhieren a (1) y (2) como la de Davidson. Con respecto a lo segundo, se puede ser instrumentalista, esto es, sostener que adscribir creencias forma parte de adoptar una actitud hacia ciertos sistemas de manera de tratarlos como intencionales sin postular que tengan estados internos intencionales y, al mismo tiempo, externalista porque al momento de adscribir deseos y creencias para dar cuenta de la conducta de esos sistemas se toma en cuenta el entorno que los rodea.

Por otro lado, si de lo que se trata es de dar condiciones suficientes naturales para que un estado físico tenga un contenido es porque se cree que dichas condiciones pueden especificar un contenido *determinado* de un estado mental particular, de manera que ambas teorías sostienen:

(7) *Determinación del contenido*: siempre hay un hecho decisivo (*fact of the matter*) con respecto a cuál es el contenido de un cierto estado intencional.

La idea es que si no hubiera un hecho objetivo en el mundo (y que pueda especificarse también de manera objetiva) que determine acerca de qué es el contenido de un estado mental, entonces las condiciones naturales no serían suficientes para determinarlo. En la TDA, el contenido está determinado por relaciones asimétricas entre contrafácticos. En la teleosemántica el contenido está fijado por las relaciones de correspondencia y las condiciones Normales. De manera que (7) se formula de estas dos formas:

---

<sup>28</sup> Hay diferentes maneras de formular las tesis externalista e internalista. En términos de superveniencia, la tesis externalista sostiene que el contenido superviene sobre propiedades físicas relacionales de (o extrínsecas a) la persona, mientras que para la internalista, el contenido superviene sobre propiedades físicas intrínsecas al individuo.

(7F) El contenido queda fijado por relaciones nomológicas asimétricas entre instanciaciones de representaciones e instanciaciones de propiedades en el mundo.

(7M) El contenido queda fijado por relaciones de proyección entre una representación y lo representado en condiciones Normales.

Cualquier teoría del contenido debe enfrentarse a (lo que en la literatura sobre el tema suelen llamarse) problemas de indeterminación como el 'problema de la disyunción': habría casos en los que el contenido de un estado mental parece ser una disyunción de contenidos ya sea porque hay varias propiedades en el mundo que podrían causar una misma representación o hay varias condiciones Normales bajo las que un sistema ejecutaría su función propia (como el caso que hemos ilustrado antes acerca del problema de determinar el contenido de VACA dado que éste es causado tanto por vacas o caballos en noches oscuras). Esto traería como consecuencia que no habría lugar para las representaciones falsas o erróneas. La TDA y la teleosemántica consideran que (7F) y (7M) respectivamente ofrecen una solución a este problema sin que intervengan factores interpretativos que hagan zozobrar el proyecto de naturalización. Sin embargo, hay quienes creen que ninguna de estas propuestas logra solucionar el problema, esto se verá en el capítulo siguiente.

La tesis (7) parece estar estrechamente relacionada con:

(8) *Atomismo del contenido*: la identidad del contenido de un estado intencional no depende de los estados intencionales asociados con ese estado intencional.<sup>29</sup>

---

<sup>29</sup> Hay otras maneras de formular la distinción entre el atomismo (o localismo) y el holismo del contenido que no son exactamente equivalentes: se puede apelar a conexiones inferenciales (Guttenplan 1994) o a propiedades (Fodor y Lepore 1992, cap. 1).

Por otro lado, entiendo que la polémica atomismo-holismo es una polémica acerca de la individuación de los contenidos y por ende, se inscribe en el ámbito epistemológico. En cambio, la polémica externalismo-internalismo es una polémica acerca de la determinación de los contenidos, y por ende, se inscribe en el ámbito metafísico. Sin embargo, todo depende de cómo se formulen las tesis. Por ejemplo, si la tesis atomista se formula en términos de propiedades atómicas, se convierte en una tesis metafísica, y si la tesis externalista se formula en términos de individuación, se

Esta tesis afirma que la identidad del contenido de una creencia no depende de las creencias asociadas, y en este sentido se opone al holismo de lo mental que sostiene que la identidad del contenido de una creencia depende de los contenidos del sistema completo de nuestras creencias (o, al menos, de un subconjunto de creencias, en la versión molecularista). Un holismo así tendría consecuencias terribles para el PNC. Primero, amenazaría la tesis de la determinación del contenido (tesis (7)), porque si el contenido de una creencia depende del resto (o parte) de las creencias, sería muy difícil la individuación de un contenido determinado. Segundo, resultaría difícil que dos personas compartan una creencia determinada porque en cada una podría estar asociada con distintas creencias y por ende, el contenido sería distinto (ni siquiera un mismo agente podría tener la misma creencia en momentos distintos). Y si no se puede establecer que, al menos, dos personas comparten creencias entonces no habría leyes (o generalizaciones) intencionales que den cuenta del comportamiento de los individuos en función de creencias compartidas y por ende, no se podría obtener una psicología científica (ver más adelante la tesis (10)). Y, tercero, el proyecto mismo de naturalización se vería amenazado porque si interviene algún aspecto epistémico en la determinación del contenido, entonces ya no es posible una explicación de la intencionalidad sólo en términos naturalistas no-semánticos/intencionales, y por ende las tesis (1) y (2) caen.

En la teoría de Fodor, el contenido sólo depende de la denotación, por eso ningún contenido depende de otro. En la teoría de Millikan, si bien la determinación del contenido no depende sólo de la denotación, puesto que se determina por medio de una función semántica de proyección hacia el mundo (Millikan 1991), esta última no lleva al holismo ni a la indeterminación. En las teorías isomorfistas la correspondencia se da entre dos estructuras: la representacional y el estado de cosas en el mundo. Los elementos de la estructura representacional representan biunívocamente elementos de una estructura en el mundo porque hay isomorfismo entre las estructuras. En este sentido, los elementos (las representaciones o los

---

convierte en una tesis epistemológica. El sentido en el que se las utiliza en cada oportunidad quedará claro por el contexto.

conceptos) de las representaciones articuladas (las creencias) son sensibles al contexto en el que aparecen, de manera que no tienen contenido independientemente de la representación articulada. Sin embargo, el contenido de las representaciones articuladas no depende del de las otras, es decir, el contenido de una creencia no depende del contenido de otras creencias del sistema, y es en este sentido que la teoría de Millikan no es holista.<sup>30</sup>

En otras palabras, lo que tiene contenido por participar de las relaciones de correspondencia con el mundo son las representaciones complejas, pero el contenido de estas representaciones sólo depende de sus relaciones de isomorfismo con estados de cosas en el mundo. Así, aunque los elementos constituyentes de las creencias no tienen contenido con independencia de la creencia en la que participan (cosa que no ocurre en la semántica informacional porque los constituyentes de las creencias tienen contenido con independencia del contexto porque se conectan nómicamente con propiedades en el mundo), las creencias son atómicas porque su contenido no depende del de otras creencias, de manera que el sistema de creencias es atomista.<sup>31</sup> En términos más específicos, el atomismo es una propiedad del sistema representacional y no de las representaciones.

Sin embargo, Millikan (1993a) sostiene que el atomismo sería una propiedad de las representaciones simples, en el caso de las que participan de las oraciones mentales (si es que hubiese algo como las oraciones mentales). La composicionalidad es una forma de proyección en la medida en que es una operación (en sentido matemático. Otras operaciones serían, por ejemplo, la sustitución y la introducción de la negación) que permite transformaciones en las representaciones. Y, para que haya composicionalidad se requiere que las representaciones hagan la misma contribución semántica en las distintas combinaciones en las que aparecen. Esto supone que las propiedades semánticas de las representaciones simbólicas son independientes del contexto y si lo son, es porque no son propiedades holistas o molecularistas.

---

<sup>30</sup> Cf. Gomila (1995, p. 108) quien afirma que la teoría telcofuncional de Millikan es claramente atomista en el sentido de que "el contenido de un estado mental no depende del contenido de otros estados mentales".

<sup>31</sup> Esta es la diferencia entre sensibilidad al contexto y holismo. Ver Cummins (1996), pp. 76-8.

Por otro lado, en la teoría del isomorfismo, las representaciones son figuras (*pictures*) en sentido abstracto y están articuladas, de manera que representan estados de cosas.<sup>32</sup> Una objeción habitual a este tipo de teorías es que hay infinitas relaciones de correspondencia, esto es, muchas correspondencias uno a uno entre dos estructuras. Millikan advierte esta dificultad pero sostiene que lo único que afirma su teoría es que *se supone* que las creencias se proyectan en estados de hecho en el mundo de acuerdo con ciertas reglas de proyección. Este *se supone* se explica en términos de condiciones Normales, de manera que sólo las creencias que son verdaderas lo son en virtud de que se ha establecido, de hecho, cierta correspondencia.

Nos queda por ver las dos últimas tesis del PNC. La idea de que la conducta de los individuos es producto de sus deseos y creencias es lo que hemos visto en el primer capítulo como PSC. Fodor cree que la PSC no sólo tiene que ser el punto de partida de la psicología sino que la psicología cognitiva no es más que la formulación científica de la PSC. Desde la PSC se suele hacer generalizaciones intencionales para dar cuenta de la conducta de los individuos. Por ejemplo, si se está incendiando un edificio se puede describir el comportamiento de los individuos que se encuentran en él diciendo que dada su creencia de que el edificio se incendia y su deseo de salir, intentarán por todos los medios salir del edificio. Si estas generalizaciones pueden convertirse en leyes entonces, la psicología tendría las leyes ('intencionales' porque relacionan contenidos y 'causales' porque lo que subsumen son las propiedades causales de los contenidos) que son necesarias en toda teoría que aspire a ser científica.

Si las leyes causales intencionales no fueran posibles, entonces no se podría incluir a la psicología dentro de la visión científica del mundo. Este es un aspecto de la empresa fodoriana de obtención de una psicología científica. El otro aspecto, igualmente necesario, es la elaboración de leyes psicofísicas que son las que determinan el contenido amplio que aparece en las leyes intencionales, esto es, la elaboración de una semántica que relacione los contenidos de las representaciones

---

<sup>32</sup> En el sentido de Wittgenstein (1921). Cf. Millikan (1990), pp. 158-9.

con rasgos del mundo. Así, la semántica del contenido mental juega un rol fundamental en la configuración del contenido que participa en las leyes causales intencionales de la futura psicología científica.

Millikan (1986) sostiene que la PSC puede ser el punto de partida para la ciencia cognitiva si se la entiende en términos de una teoría que pretende explicar el comportamiento de los organismos postulando entidades con funciones propias, y no entidades que obedecen leyes. Así, la psicología no sería científica en el sentido de constar de leyes sino que tan sólo se podrían hacer generalizaciones. Esto es así porque las funciones propias no se determinan por disposiciones efectivas de los sistemas cognitivos sino por su historia evolutiva. Las funciones propias no siempre se realizan porque dependen de mecanismos que pueden funcionar defectuosamente y de condiciones externas Normales que pueden fallar. De manera que las funciones propias no se corresponden con leyes causales estrictas (válidas para todos los casos) ni estadísticas.

La diferencia de criterios entre las posturas de Fodor y Millikan radica en modelos distintos para la explicación psicológica. El modelo fodoriano es el de subsunción bajo leyes causales mientras que Millikan (1999b) apela a las explicaciones históricas. La psicología (como la biología y otras ciencias) es histórica en tanto que sus clases naturales se definen en virtud de relaciones históricas entre sus miembros. Estas relaciones están dadas por pertenecer a la misma familia reproductiva (o de copia) y compartir el mismo entorno. Las inducciones de un miembro a otro de la clase o del comportamiento futuro de un miembro funcionan porque hay razones históricas (/biológicas). Sin embargo, los procesos de copia y el entorno histórico no siempre son los adecuados de manera que no hay que esperar explicaciones psicológicas predictivas como las causales (Millikan 1993b). De todas formas, habría explicaciones de cómo un sistema cognitivo ejecuta Normalmente las funciones y, en cierto grado falible, se podría predecir su comportamiento. Así, las explicaciones no serían según el modelo de cobertura legal-causal (haciendo referencia a condiciones iniciales y leyes), sino que serían como las históricas, esto es, retrospectivas, tratando de explicar cómo sucedió un episodio haciendo alusión a los pasos de una secuencia.



Este argumento de Millikan no parece ser una razón de peso para la inexistencia de leyes, puesto que Fodor tampoco cree que las leyes de la psicología científica serán leyes estrictas (como las de la física) sino *ceteris paribus*, esto es, leyes que se cumplen manteniéndose ciertas condiciones iguales y en este sentido, admiten excepciones. Así, se podría pensar que los ítems que tienen funciones, las realizan conforme a leyes *ceteris paribus*. Sin embargo, Millikan (1993b) sostiene que no son posibles las leyes *ceteris paribus* que cubran todos los casos que tienen una cierta función. Estas leyes requieren que ciertas condiciones no especificadas se mantengan igual de caso en caso, pero hay ítems con funciones propias defectuosas (por ejemplo, un corazón enfermo), hay ítems cuyas funciones rara vez se realizan (el caso de los espermatozoides), y la misma función puede operar según distintos principios que requieren de distintas condiciones (por ejemplo, las células cerebrales requieren oxígeno mientras que los chips de una computadora requieren corriente eléctrica). Así, habría que agregar diferentes condiciones para cada caso, de manera que las leyes *ceteris paribus* se volverían vacías. Lo que sí estaría dispuesta a admitir Millikan es la posibilidad de leyes *ceteris paribus* no-causales para distintas especies (humanos, marcianos, y demás). De modo que, en última instancia tanto para la TDA como para la teleosemántica serían posibles las leyes *ceteris paribus* para los humanos, la diferencia radica en que para la concepción fodoriana éstas serían causales y para Millikan no.

No es difícil advertir, en esta concepción de la psicología científica, el lugar que ocupa una teoría del contenido. Puesto que las generalizaciones que dan cuenta de cómo los mecanismos cognitivos realizan sus funciones incluyen las condiciones Normales, esto es, los contenidos, es necesario desarrollar una semántica del contenido mental. Al igual que en la concepción fodoriana, la semántica del contenido mental juega un rol importante en la psicología cognitiva. Ahora bien, más allá del desacuerdo entre ambas teorías acerca de cómo entender a la PSC y en consecuencia a la psicología cognitiva, está claro que ambas consideran que la PSC es el punto de partida para la obtención de una psicología científica, por ende, acuerdan en:

(9) *Psicología de sentido común*: el punto de partida de la psicología cognitiva es la psicología de sentido común.

(10) *Psicología científica*: uno de los objetivos (explícitos o implícitos) para el desarrollo de una semántica del contenido mental es lograr una psicología científica.

Hasta aquí hemos visto que el PNC está comprometido con estas diez tesis.<sup>33</sup> También vimos algunas de las relaciones entre ellas (que se desarrollarán con más detalle en el capítulo siguiente) y los diferentes enfoques que adoptan ambas teorías respecto de algunas de ellas. Quisiera, para concluir este capítulo, abordar una cuestión que a veces ha llevado a algunos filósofos a pensar que la teleosemántica a diferencia de la semántica informacional puede dar una respuesta al problema de la relación entre el ámbito causal y el normativo, incorporando un elemento normativo en la explicación naturalista de la intencionalidad.

Los estados intencionales parecen tener propiedades normativas, esto es, evaluativas, prescriptivas. Las creencias pueden ser correctas o incorrectas, apropiadas o inapropiadas, aceptables o inaceptables, verdaderas o falsas. Para evaluar la corrección de una creencia tiene que haber un estándar con el cual comparar su corrección, algo que diga cuál *debe* ser la creencia correcta en un caso particular.<sup>34</sup> Así, hay teorías del contenido que consideran que hay normas que gobiernan la atribución de creencias, y en este sentido la normatividad es constitutiva del contenido intencional. Ejemplos de estas teorías son los enfoques dummiettianos o davidsonianos, en los que el contenido depende de las condiciones en las que uno está autorizado a tener una creencia o depende, entre otras cosas, de cuando es razonable

---

<sup>33</sup> Por supuesto que es discutible que algunas de las tesis tengan que formar parte necesariamente del PNC (por ejemplo, otras teorías naturalistas científicas que no consideré podrían no adscribir a alguna de ellas). Sin embargo, por un lado, no se puede considerar a una teoría parte del PNC si rechaza alguna de sus tesis centrales (quizás (1)-(5), (7)-(8), y (10)), y, por otro lado, las tesis restantes son las que se pueden encontrar en las semánticas en las que el PNC se basa.

<sup>34</sup> El paralelo semántico sería el siguiente: "El punto crucial de la interpretación que hace Kripke de Wittgenstein puede ponerse así. Es parte de la esencia de que una expresión signifique de una cierta manera, que el que signifique de esa manera determine cómo tendría que usarse la expresión si se la usa correctamente... Cualquier candidato propuesto para ser la propiedad en virtud de la

formar una creencia (Cf. Papineau 1999). Así, las normas constitutivas del contenido de los estados intencionales son prescriptivas en la medida en que dadas ciertas consideraciones relativas al individuo (de racionalidad, su lugar en el mundo, su sistema cognitivo, sus necesidades, su biografía, y demás), las creencias que posee son las que *debería* tener.

Por supuesto que ni la semántica informacional ni la teleosemántica sostienen que el fenómeno de la intencionalidad es normativo en este sentido. Ninguna teoría naturalista del contenido aceptaría que hay normas evaluativas/prescriptivas constitutivas del contenido porque el rechazo de esto es justamente lo que las hace ser naturalistas. Si se aceptara este sentido de normatividad, las condiciones naturales para que un individuo tuviera creencias no serían suficientes y con ello se pondría en peligro el proyecto naturalista puesto que tener una creencia no consistiría solamente en que su contenido esté conectado naturalmente con el mundo sino en aplicarla con algún estándar de corrección (algo así como estar justificada). Las teorías naturalistas no niegan ni la racionalidad de los agentes cognitivos ni la dimensión normativa de la atribución de creencias (en función de principios de racionalidad o normas de corrección esenciales a los significados), sólo que se proponen dar cuenta de la racionalidad en términos puramente no intencionales y en ese sentido, en la determinación del contenido no interviene ningún elemento normativo constitutivo.<sup>35</sup>

Sin embargo, Millikan (1984, 1991) sostiene que la teleosemántica rescata un aspecto normativo dado que la noción de función propia es normativa. Puesto que la función se identifica con independencia de lo que el organismo hace efectivamente (esto es, no es una noción causal ni disposicional), hay un aspecto relacionado con lo que la función *fue diseñada para hacer* o *lo que se supone que debe hacer* que actuaría de estándar a partir del cual poder evaluar si el sistema está haciendo o no aquello para lo que fue diseñado. Se puede decir, por ejemplo, que el corazón no

---

cual una expresión tiene significado debe ser tal que funde la 'normatividad' del significado" (Boghossian, 1989, pp. 83-84, citado en Fodor 1990a, nota 35, p. 135).

<sup>35</sup> La respuesta de un naturalista a posturas como la de la nota anterior es que para aplicar correctamente el término "mesa" no se requiere de nada más que aplicarlo a una cosa que cae bajo su extensión, así "una vez que se ha dicho lo que hace que las mesas sean la extensión de "mesas", seguramente no hay ninguna cuestión *adicional* acerca de por qué es *correcto* aplicar "mesa" a una mesa" (Fodor 1990a, nota 35, p. 135).

funciona bien porque no bombea sangre al cuerpo en virtud de que el corazón *debe* bombear sangre. Esto no perjudica la naturalización del contenido porque éste no depende de (no queda determinado por) la función propia sino de (por) las condiciones naturales Normales. Así, las representaciones tienen un aspecto natural que es el que fija su contenido y heredan la normatividad del diseño de los sistemas en los que intervienen.

Algunos han visto en estrategias como esta una manera de resolver el conflicto entre naturalismo y normatividad. Así, McGuinn dice con respecto al mal funcionamiento de la danza de las abejas por un defecto en una de ellas que:

Al no danzar como se supone biológicamente que debería hacerlo está violando una norma de las abejas de señalización del néctar. Así, aquí hay lugar para una distinción entre *es* y *debe*, y por ende, los rudimentos de una noción normativa de contenido. El punto básico es simplemente que la noción de función propia, siendo normativa, provee de una base para el tipo de distinción entre lo correcto y lo incorrecto que esperamos de una teoría del contenido. La normatividad semántica surge de la normatividad funcional. (1989, pp. 160-61)

Sin embargo, la cita de McGuinn sólo muestra un desliz apresurado entre un sentido de normatividad que no es prescriptivo hacia otro que lo sería. Las funciones propias no son normas biológicas en ningún sentido prescriptivo, no dicen lo que un sistema *debe* (prescriptivamente) hacer. Las normas biológicas dicen que un sistema opera de cierta forma por haber sido diseñado así, pero según Papineau:

Es un error grosero, y más aún, peligroso, inferir de la premisa de que X ha sido biológicamente diseñado para Y, que en algún sentido X *debe* Y. Podría decirse que mis nudillos han sido diseñados biológicamente para pegarle a la gente, pero no se sigue en ningún sentido que deba usarlos. Del mismo modo, una cantidad de rasgos humanos masculinos han sido diseñados indudablemente para fomentar la rapacidad

sexual, pero no se sigue que los hombres deban ser depredadores sexuales. (1999, n.5, p. 18)<sup>36</sup>

Millikan (2000, 2002) acuerda en que las funciones propias no son normas biológicas en ningún sentido evaluativo o prescriptivo, pero cree que lo normativo no se agota en esos sentidos. La noción de función propia sigue siendo normativa en el sentido de que hay términos normativos que se los utiliza para señalar una medida a partir de la cual son posibles las desviaciones efectivas. En este sentido, un número promedio así como toda regularidad es una norma. Aunque las funciones propias, tal como hemos visto, no corresponden a correlaciones estadísticas ni regularidades sirven para señalar una medida a partir de la cual se miden los casos de mal funcionamiento. Si este es el sentido de norma biológica que se utiliza en la teleosemántica difícilmente pueda servir como punto de partida para el surgimiento de la normatividad, tal como se la suele entender, en sentido prescriptivo. Pero está bien que esto sea así porque, tal como se ha dicho, si interviene algún elemento prescriptivo en la determinación de los contenidos, las condiciones naturales para que un estado físico tenga contenido no serían suficientes.

En este capítulo he presentado el proyecto de naturalización científicista de los contenidos intencionales. En el capítulo siguiente ofreceré algunas estrategias de ataque a este proyecto y optaré por discutir, desde una concepción diferente de la relación individuo-mundo, cada una de las diez tesis desarrolladas, con el objetivo de esbozar los lineamientos de un enfoque naturalista no-científicista del fenómeno de la intencionalidad.

---

<sup>36</sup> La paginación de la cita es del artículo tal como figura on-line. Un argumento similar se encuentra en Woodfield (1990).

## 4

## Naturalismo sin semántica intencional

En el capítulo anterior hemos visto las tesis básicas del PNC, algunas de las relaciones entre ellas y las motivaciones para el desarrollo de una semántica del contenido de los estados intencionales. Actualmente, el PNC parece ser la única opción naturalista posible para dar cuenta del fenómeno de la intencionalidad. Y, esta única opción sostiene que el ámbito representacional, y consecuentemente el ámbito de incumbencia de la psicología cognitiva, es el terreno adecuado para dilucidar la cuestión de la intencionalidad. Sin embargo, es posible delinear las bases de un Proyecto Naturalista Amplio (PNA) en el que no se requiera de una semántica intencional para las representaciones mentales. Mi intención es mostrar, en este capítulo, que la explicación del fenómeno de la intencionalidad no requiere del PNC y, en el capítulo siguiente, que los estados cognitivos que figuran en las teorías psicológicas tampoco requieren del PNC.

La primera parte de este capítulo está dedicada a la evaluación crítica del PNC. Presento cuatro estrategias posibles para objetar este proyecto. En primer lugar, dar argumentos *a priori* que muestren que un proyecto así no sería viable. Segundo, formular críticas a la semántica informacional y la teleosemántica, de manera de mostrar que no logran conectar los estados mentales con algo en el mundo y por ende, que el proyecto no ha dado resultados. Tercero, discutir las tesis básicas del PNC y evaluar cuáles se está dispuesto a aceptar y cuáles no. Y, por último, oponerle otra concepción de la relación entre el individuo y el mundo, de manera de mostrar que un proyecto de semántica para los contenidos mentales es superfluo.

Las dos primeras han sido las estrategias generalmente adoptadas y no parecen haber socavado el PNC, de manera que realizo un breve recorrido por las objeciones más habituales y opto por las dos últimas estrategias a las que considero

solidarias para el desarrollo de un naturalismo amplio. Así, en la segunda parte del capítulo desarrollo las tesis básicas del PNA. Las tesis del PNA forman parte de un proyecto programático a partir del cual podría darse cuenta del fenómeno de la intencionalidad y del tipo de contenido que interviene en la explicación psicológica. La idea que subyace al PNA es que no debe asimilarse intencionalidad a representación tal como la filosofía de la psicología hizo desde sus comienzos (ver capítulos 1 y 2). Así, sería posible dar cuenta de la intencionalidad reubicándola en el ámbito de la comunicación, y dejar el ámbito de la representación para cuestiones psicológicas. Es en este sentido que el PNA es el marco adecuado para el futuro desarrollo de una teoría de la atribución intencional y de una teoría de las representaciones que satisfaga los requisitos enunciados en el capítulo 2. Tal como mencioné en ese capítulo, mi intención es tan sólo plantear los lineamientos generales o el marco conceptual a partir del cual sería viable el desarrollo de ambas teorías.

## 1. Críticas al naturalismo científicista

A continuación se formulan en conjunto las tesis del PNC para facilitar su discusión y comparación con el PNA que se delinearán en la segunda parte del capítulo (nótese que las tres primeras tesis son sobre el naturalismo, las cinco siguientes sobre la naturaleza de los estados intencionales, y las dos últimas sobre la psicología científica):

(1) *Naturalismo*: las condiciones suficientes para que un estado físico tenga contenido se especifican en términos no-semánticos y no-intencionales.

(2) *Naturalismo científicista*: los términos no-intencionales usados en (1) son términos que refieren a propiedades (eventos, procesos, relaciones, y demás) que las ciencias naturales aceptan en sus teorías.

(3) *Naturalismo metafilosófico*: la filosofía de la mente se desarrolla en conformidad con, o hace uso de, o hace explícito el aparato metafísico de, la ciencia corriente de manera de integrar la mente en el mundo natural científico.

(4) *Realismo intencional*: hay estados internos con contenido intencional que es causalmente explicativo de la conducta del organismo y es semánticamente evaluable.

(5) *Representacionalismo intencional*: las representaciones mentales son las portadoras de la intencionalidad de los estados intencionales.

(6) *Externalismo*: el contenido de los estados intencionales de una persona está determinado por sus relaciones con el mundo

(7) *Determinación del contenido*: siempre hay un hecho decisivo con respecto a cuál es el contenido de un cierto estado intencional.

(8) *Atomismo del contenido*: la identidad del contenido de un estado intencional no depende de los estados intencionales asociados con ese estado intencional.

(9) *Psicología de sentido común*: el punto de partida de la psicología cognitiva es la psicología de sentido común.

(10) *Psicología científica*: uno de los objetivos (explícitos o implícitos) para el desarrollo de una semántica del contenido mental es lograr una psicología científica.

Si uno quisiera oponerse a este tipo de proyecto, sin dejar de ser naturalista en un sentido más amplio que el naturalismo científicista, podría adoptar distintas



estrategias. Una de ellas consistiría en dar argumentos *a priori* con la intención de que sean contundentes y demolidores acerca de por qué un proyecto así no sería viable. Una segunda estrategia podría ser la de formular críticas a la semántica informacional y la teleosemántica, de manera de mostrar que no logran conectar los estados mentales con algo en el mundo y por ende, que el proyecto no ha dado resultados. La diferencia con la primera estrategia es que aquí no se esgrime la imposibilidad metafísica del proyecto sino problemas, que pueden ser metafísicos o de otra índole, de las teorías en oferta para llevarlo a cabo.

Ambas teorías han recibido críticas y en particular tienen dificultades con la cuestión de la determinación del contenido. Por supuesto, esto no significa que el proyecto no pueda tener éxito en el futuro pero una acumulación significativa de problemas sería una razón de cierto peso para abandonarlo. Por otro lado, la crítica interna, en tanto crítica sólo se atiene al aspecto negativo sin ofrecer una propuesta concreta de un proyecto naturalista no cientificista, y en tanto interna, no suele socavar los presupuestos básicos de una teoría sino que los da por sentado, es algo así como una disputa *dentro* de la familia cuando de lo que se trata es de discutir las mismas bases del proyecto, esto es una disputa *entre* familias.

Una tercera estrategia sería optar por discutir las tesis básicas del PNC y evaluar cuáles se está dispuesto a aceptar y cuáles no y ver qué tipo de proyecto naturalista quedaría. Una cuarta estrategia sería directamente oponerle otra concepción de la relación entre el individuo y el mundo de manera de mostrar que un proyecto de semántica para los contenidos mentales es superfluo. Considero que estas dos últimas estrategias son solidarias. En lo que sigue me ocuparé de reconstruir, en el marco de la primera y segunda estrategias, algunas críticas que el PNC ha recibido, y presentaré, en el marco de la tercera y cuarta estrategias, algunas líneas para una opción naturalista diferente al PNC.

### *1.1. Primera estrategia: argumentos a priori*

Un argumento *a priori* que puede utilizarse contra la posibilidad de la naturalización del contenido es el de la imposibilidad de la existencia de leyes psicofísicas. Davidson (1970) sostiene que los términos mentales y físicos pertenecen a ámbitos distintos regidos por principios constitutivos totalmente diferentes. Los principios de racionalidad son constitutivos del ámbito de lo mental en el sentido de que constriñen la atribución de estados mentales. Dado que lo mental es holista, atribuir un estado mental requiere atribuir coherencia, racionalidad y consistencia al sistema de creencias del individuo al que se le atribuye el estado mental.<sup>1</sup> En cambio, estos principios no rigen en el ámbito de lo físico. Estos ámbitos se verían seriamente comprometidos en el caso de que se conectasen por medio de leyes porque transmitirían sus rasgos al otro ámbito. De manera que, o bien la racionalidad quedaría excluida, lo cual sería inadmisible o bien se introduciría la racionalidad en el ámbito de lo físico, lo cual también sería inadmisible.<sup>2</sup>

Si esto es así y no sería posible la obtención de leyes psicofísicas, entonces, las condiciones naturales esgrimidas por Fodor, que dependen de la existencia de leyes psicofísicas que conecten propiedades del mundo con propiedades mentales, no se obtendrían.<sup>3</sup> Aún si se acuerda con este argumento *a priori* (Davidson mismo considera que sólo puede dar razones, y no pruebas, de que no haya leyes psicofísicas), una solución podría ser optar por las condiciones naturales que ofrece Millikan, las cuales, tal como hemos visto en el capítulo anterior, no parecen depender de la existencia de leyes psicofísicas. El argumento *a priori* de Davidson no va en contra de la obtención de generalizaciones psicofísicas verdaderas en psicología sino de que pueda haber leyes (estrictas), a la manera en que el argumento empírico

---

<sup>1</sup> Este es el sentido de normatividad que hemos visto en el capítulo anterior que sería incompatible con el PNC.

<sup>2</sup> Para una reconstrucción más detallada del argumento de Davidson, véase Kim (1993).

<sup>3</sup> Recuérdese que una ley, como por ejemplo, 'las vacas causan *vacas* o pensamientos acerca de vacas' es una ley psicofísica porque relaciona algo del mundo físico con algo mental. El argumento de Davidson fue pensado para la relación psicofísica mente-cuerpo, su extensión a la relación psicofísica mente-mundo es directa dado que ambas relaciones se dan entre el ámbito mental y el físico.

de Millikan va en contra de que haya leyes (estrictas) en psicología, sosteniendo que lo único que se pueden obtener son generalizaciones.<sup>4</sup>

Sin embargo, si se acepta la conclusión fuerte del argumento, esto es, que lo que muestra es la imposibilidad de la reducción de los conceptos mentales a los físicos, en otras palabras, la imposibilidad de la naturalización del contenido, siempre está la posibilidad abierta de que puesto que, desde un punto de vista fáctico, no hay pruebas de que esto no pueda ser así y puesto que el proyecto naturalista se plantea como un desafío empírico, entonces, la cuestión de su inviabilidad no se dirime en un terreno puramente *a priori*. Tal como afirma Fodor (1987, p. 180), lo único que se necesita probar para ganarle a quien considera que hay una razón metafísica *a priori* para sostener que la semántica no puede naturalizarse, es que hay condiciones naturales suficientes para que "P" signifique *p*. De manera que la estrategia a seguir sería la de rechazar los planteos *a priori* porque la cuestión de la naturalización de los contenidos se dirime en el ámbito empírico.

### 1.2. Segunda estrategia: críticas internas

La TDA ha recibido críticas que van desde cuestiones puntuales de la teoría a otras más globales o radicales que atacan el proyecto mismo de naturalización. Las más puntuales objetan la condición de dependencia asimétrica (condición 2 del §1 del capítulo 3), de manera que las condiciones naturales establecidas por Fodor no serían suficientes para determinar el contenido. Aquí se verán brevemente las objeciones de Adams y Aizawa (1994), Papineau (1994) y Millikan (1991). Por otro lado, las críticas globales ponen en duda que la semántica de Fodor sea naturalista, en particular, es Putnam (1992) quien plantea esta objeción. La diferencia entre ambos tipos de crítica radica en que las puntuales acuerdan con el PNC, a diferencia de las globales que ponen en duda la concepción del naturalismo que sustenta a la TDA (en particular, las tesis (1) y (2) del PNC).

---

<sup>4</sup> Aunque sí habría un conjunto de leyes no-causales *ceteris paribus* para los humanos, otro conjunto para los animales, y así en más. Ver §2 del capítulo 3.

Adams y Aizawa (1994) muestran, por un lado, que puede haber una ley que conecte "X" con Y que no dependa asimétricamente de la ley que conecta "X" con X, y por otro lado, que puede haber una dependencia simétrica entre las leyes que conectan "X" con Y, y "X" con X. Con respecto a lo primero, podría darse el caso de que haya causas patológicas de "X", como drogas alucinógenas o un tumor cerebral o microelectrodos, de manera que, por ejemplo, la ley "agua"-microelectrodos ("X"-Y) no depende sincrónicamente de la ley "agua"-H<sub>2</sub>O ("X"-X). Esto es, el poder de los microelectrodos para causar "agua" no depende sincrónicamente del poder de H<sub>2</sub>O para causar "agua". Con respecto a lo segundo, las leyes que conectan "agua"- H<sub>2</sub>O y "agua"-XYZ son simétricamente dependientes para alguien que no distingue entre H<sub>2</sub>O y XYZ, en este sentido, si se rompe cualquiera de esas conexiones se quiebra la otra. Si estas objeciones son adecuadas, la condición de dependencia asimétrica no se cumpliría y por ende, el contenido quedaría indeterminado.

Por otra parte, tanto Papineau (1994) como Millikan (1991) acusan a la TDA de circularidad, ya sea en la forma de una petición de principio o un regreso. El argumento consiste en que la asimetría entre, por ejemplo, "las vacas causan pensamientos acerca de vacas" y "los caballos en noches oscuras causan pensamientos acerca de vacas", descansa en que los pensamientos acerca de las vacas ya son acerca de ellas. Así, para Papineau, la teoría supone lo que quiere explicar, de manera que la asimetría es más bien una consecuencia que el fundamento del contenido de los pensamientos acerca de las vacas. Y, para Millikan, la dependencia no es causal sino teleológica porque no hay manera de determinar el contenido sin que se genere un regreso de las relaciones causales. Si estas objeciones son adecuadas, entonces la TDA no logra conectar lo mental con el mundo.

Putnam (1992) objeta a la teoría de Fodor que no logra naturalizar el contenido porque apela a las nociones de "causa" y "contrafáctico" que son intencionales o normativas. La noción de "causalidad" (del sentido común, que es, según Putnam, a la que Fodor alude) tiene un aspecto cognitivo relacionado con lo que un individuo conoce acerca de los elementos que intervienen en una relación causal, es decir que la noción es sensible al contexto y los intereses de los individuos. Con respecto a la noción de "contrafáctico", interviene un elemento relacionado con

las intenciones del hablante que lo emite y/o valora, este elemento consiste en la razonabilidad de la descripción de las condiciones relevantes que harían verdadero a un contrafáctico (esto es, las condiciones relevantes en las cuales no se daría el caso que el antecedente fuera verdadero y el consecuente falso). En este sentido, hablar de razonabilidad o irrazonabilidad implica un juicio normativo.

Cabe recordar que, en la TDA, las condiciones naturales suficientes para que un estado físico tenga un contenido determinado se expresan en términos de relaciones causales y expresiones que denotan objetos intensionales como leyes, contrafácticos y propiedades. La defensa de la TDA podría adoptar la siguiente forma: por un lado, en relación con la noción de causalidad, podría argüir que ésta es la que usan las ciencias naturales (ver nota 7 del capítulo 3), y por el otro, el proyecto de naturalización (cualquiera sea la teoría sustancial) no excluye la utilización de nociones intensionales como las mencionadas o las teleológicas (ver §1 del capítulo 3). Una razón en favor de esto último es que el fisicalismo no excluye un realismo acerca de las entidades abstractas (Fodor 1990b, p. 132, n. 6). De modo que, dado que las ciencias naturales hacen uso de estas nociones (“causalidad” y “contrafácticos”), y puesto que describen el mundo de una manera no intencional (en términos aceptados por la ontología fisicalista), entonces estas nociones no serían problemáticas.

Con respecto a la teleosemántica, las críticas también van desde aspectos más puntuales como problemas en la determinación del contenido a cuestiones más generales que objetan el enfoque evolutivo, por ejemplo, la objeción de que no todo producto de la evolución natural tiene que tener una explicación en términos evolutivos. Aquí se verán brevemente las objeciones de Fodor (1990a, 1991) y Gomila Benejam (1995) con relación a las del primer tipo y la de Peacocke (1995) con respecto a las del segundo tipo. En este caso, ambas críticas no socavan el PNC sino que van dirigidas específicamente contra el enfoque teleológico como la propuesta adecuada para satisfacer las tesis del PNC.

El ejemplo más utilizado para explicitar el problema de la indeterminación del contenido como consecuencia del problema de la disyunción, es el de la alimentación de las ranas. Si se quiere explicitar el contenido del estado mental de una rana cuando

saca su lengua para alimentarse, no se sabe si representa una mosca o un punto negro pequeño en movimiento puesto que las ranas responden de la misma manera (intentando atrapar) a cualquiera de estas cosas. El contenido parece ser la disyunción de *mosca o punto negro pequeño en movimiento*, y no parece haber ningún hecho en el mundo que permita decidir cuál de estos contenidos es el correcto, con lo cual el contenido no estaría determinado. Según la teleosemántica, el cerebro de la rana representa moscas porque su función biológica de alimentarse tiene éxito cuando la condición natural son las moscas (Millikan 1991, pp. 159-60). Según Fodor (1990a), la respuesta de la teleosemántica no es adecuada porque no hay manera de establecer evolutivamente que su función es covariar con las moscas y no con los puntos negros teniendo en cuenta que esto último le sería de utilidad a la rana en un entorno donde la mayoría de los puntos negros son moscas.<sup>5</sup>

Según Millikan, la función de una creencia se relaciona no con lo que la causó sino con el beneficio que trae una vez que ya está. Por eso, hay que preguntarse por las condiciones que hacen que una creencia produzca resultados ventajosos que llevaron a su selección. Estas condiciones involucran la presencia de moscas porque es en su presencia que los mecanismos consumidores de representaciones determinan la regla de correspondencia adecuada para que la rana se comporte de manera que sus resultados sean ventajosos, en este caso, que se alimente. Fodor (1991) objeta a esta respuesta que la teleosemántica determina la función de manera arbitraria, que las teorías del isomorfismo son semánticamente indeterminadas y que esta propuesta de solución le pasa el problema de la indeterminación a los consumidores. Si estas objeciones son adecuadas, entonces la teleosemántica no logra resolver el problema de la determinación del contenido.

Por otra parte, están las críticas relacionadas con la caracterización de las condiciones Normales (que, en última instancia, derivan en el problema de la determinación del contenido, puesto que el contenido depende de las condiciones Normales). Gomila (1995) considera que éste es un problema porque habría casos de condiciones no Normales pero favorables para el buen funcionamiento; y también habría que resolver la cuestión del conjunto de las condiciones Normales relevante

---

<sup>5</sup> Ver nota 11 del capítulo 3.

según sea el caso. En este sentido, Fodor (1987, p. 151) acusa a la teleosemántica de lo mismo que ésta acusa a la TDA, esto es, de circularidad puesto que para saber cuáles son las condiciones Normales ya hay que saber cuál es el contenido. Si estas críticas resultan correctas, entonces la teleosemántica no logra conectar los estados mentales con el mundo.

Con relación al segundo tipo de críticas, Peacocke (1995) sostiene que aceptar que desde un punto de vista biológico es deseable tener estados mentales no tiene por qué llevarnos a aceptar que la funcionalidad biológica es lo que determina el contenido. Si bien Peacocke no se explaya al respecto, esto parece tener importancia porque Millikan (por ejemplo, 1989a) suele afirmar que o bien el contenido se constituye por la función biológica o bien el hecho de tener una mente es un accidente desde el punto de vista biológico. Aceptar el segundo disyunto significa aceptar que la mente surgió como un producto contingente de la evolución (es decir, como un producto no adaptativo), de manera que no tendría una función propia originaria y por ende, no tendría sentido una explicación del contenido en términos de funciones biológicas.

En principio, y a mi juicio, este parece ser un falso dilema porque se puede aceptar que la mente no es un producto azaroso y sin embargo considerar que el enfoque adecuado para el contenido de los estados mentales no es el evolutivo. Una analogía puede ilustrar el punto. Un tipo de funcionalismo considera que los estados mentales son funcionales en el sentido de que el rol que cumplen en un sistema cognitivo depende de los inputs que los causaron y los outputs que causan como así también de sus relaciones causales con otros estados mentales. Del hecho de considerar a los estados mentales como funcionales no se sigue que el contenido de esos estados se determine de manera funcional. Es perfectamente compatible ser funcionalista en relación a los estados mentales y no-funcionalista en relación a los contenidos. Este es el caso de Fodor, quien rechaza el funcionalismo para individuar el contenido porque lleva a un holismo (y además hemos visto, más arriba y en el capítulo anterior, que considera que la misma noción de función lleva a la indeterminación del contenido).

Lo que le permite a Millikan plantear este dilema es que no distingue entre que algo sea un estado mental (*que algo tenga contenido*) y que el estado mental tenga el *contenido particular* que tiene. Mientras que para Fodor lo que hace que algo sea una representación está determinado por la organización funcional del sistema y su contenido por las causas que bajo ciertas condiciones lo determinan, para Millikan, que algo sea un estado mental y lo que fija el contenido particular de ese estado mental se realiza por el mismo mecanismo de isomorfismo y responde a la misma explicación Normal. Si siguiendo a Fodor (y la mayoría de los filósofos de la mente) se acepta esta distinción, podría pensarse que tenemos estados mentales gracias a lo que la biología dice acerca del proceso por el cual surgieron, pero eso no implica que el mismo proceso pueda dar cuenta del contenido de esos estados.

En relación a ambas teorías, las críticas del primer tipo se concentran en el problema de la indeterminación del contenido. Ninguna de ellas parece poder resolverlo sin apelar a estipulaciones arbitrarias, ya sea de las funciones para el caso de la teleosemántica o de las dependencias asimétricas entre las leyes en el caso de la TDA. Para la consecución de este proyecto naturalista es imprescindible tener una respuesta apropiada puesto que si no hay determinación, entonces, como ya se ha mencionado en el capítulo anterior, las condiciones naturales no serían suficientes para dar cuenta del contenido de los estados mentales y si esto es así entonces, no habría naturalización (ni contenido, dado que contenido es igual a contenido naturalizado). Volveré sobre esto en la segunda parte de este capítulo.

### *1.3. Tercera y cuarta estrategia: hacia un naturalismo amplio*

Hay, al menos, dos cuestiones que yacen a la base de la infructuosidad del PNC. Una de ellas es la concepción estrecha de la naturalización cuyo objetivo es la explicación de la relación representacional en términos no-intencionales. La otra es la aplicación *tout court* al ámbito de lo mental del marco conceptual de las teorías semánticas para el lenguaje natural. Ambas cuestiones están íntimamente relacionadas y hasta me atrevería a decir que la segunda es, en parte, la motivación



de la concepción naturalista estrecha. A grandes rasgos, si el problema de la intencionalidad se reduce al problema de la representación y se entiende la relación de representación en términos de la noción "técnica" de "referencia" (por noción "técnica" entiendo una noción pensada o acuñada para un propósito teórico específico, en este caso, para dar cuenta de la relación semántica entre el lenguaje natural y el mundo) es muy probable que la naturalización de lo mental pase por reducir la relación semántica de referencia a una relación no-semántica especificada en términos naturales.

Si, en cambio, se concibe el fenómeno de la intencionalidad en un ámbito distinto al de la representación, y la noción de representación no se la entiende en función de las nociones técnicas de la semántica para los lenguajes naturales (de manera de elaborar nociones teóricas propias para el ámbito de las representaciones), el proyecto de naturalización de lo mental no adoptaría una perspectiva estrecha. Dicho de otra manera, si se concibe la noción de representación como una noción psicológica técnica no-intencional (de esto me ocuparé en detalle en el capítulo 5), y se concibe el fenómeno de la intencionalidad en el ámbito de las prácticas comunicativas, la naturalización de lo mental no consiste en la búsqueda de condiciones suficientes (para que, por ejemplo, AGUA signifique, en el sentido de referir, H<sub>2</sub>O), sino que aquello que hay que concebir de manera naturalista serían nuestras prácticas comunicativas. En lo que sigue me ocuparé de estas dos cuestiones, esto es, la falta de un marco de nociones propias en el ámbito de lo mental y la concepción naturalista estrecha.

Según la semántica informacional, lo que está en el mundo determina el contenido de una representación porque hay covariaciones causales que los conectan. Según la biosemántica el contenido se determina por una relación de isomorfismo estructural entre las representaciones y lo representado. Ambos tipos de teoría no son más que las versiones refinadas de teorías semánticas para el lenguaje natural: la teoría causal de la referencia en el caso de Fodor y la teoría pictórica en el caso de Millikan. Estas teorías no son desarrollos específicos para los estados mentales, esto es, no hay un marco conceptual propio para la elucidación de los mismos. La idea es tomar teorías que se han desarrollado para el lenguaje natural y, con mayor o menor

suerte en la resolución de sus problemas, utilizarlas para dar cuenta de lo mental porque simplemente se asume que las propiedades semánticas de las representaciones mentales son análogas a las del lenguaje.

Tanto la semántica informacional como la teleosemántica sostienen que las representaciones tienen propiedades sintácticas y semánticas. Las representaciones tienen la propiedad sintáctica de ser vehículos de propiedades semánticas que son los contenidos mentales; así como las palabras o las oraciones son los vehículos de los significados. Para lo que nos interesa, más allá de que el formato representacional sea o no como un lenguaje, es decir, que sean símbolos o mapas/modelos, el sistema representacional posee las mismas propiedades semánticas que un lenguaje natural (referencia, significado, condiciones de verdad, y demás). Se asume que la relación representacional es una relación semántica a la manera de las relaciones semánticas de significación o referencia.

En este sentido, el PNC concibe la relación de representación a la manera en que lo hacen las semánticas del lenguaje natural en las que se basa. Así como en la teoría causal de la referencia, las palabras refieren a cosas en el mundo, en la TDA las representaciones (primitivas) refieren a propiedades del mundo. Y, así como en la teoría pictórica las proposiciones refieren a estados de cosas en el mundo, en la teleosemántica las representaciones (articuladas) refieren a estados de cosas en el mundo. La relación de representación se concibe a la manera en que se desarrolló la noción filosófica técnica de "referencia" como una relación entre las representaciones y algo particular en el mundo.

La cuestión es, por un lado, si la utilización de esta noción técnica puede echar luz sobre el fenómeno de la representación tal como éste es abordado en las teorías psicológicas (volveré sobre esto en el §2 y en el capítulo siguiente), y por el otro, si se avanza en algo en cuanto a la comprensión del fenómeno de la representación utilizando teorías que arrastran los mismos problemas que tuvieron para dar cuenta del lenguaje natural. Si la estrategia es tomar teorías ya establecidas (aunque refinadas y detalladas) y aplicarlas *tout court* a los estados mentales, no se ve qué es lo que impediría en este nuevo terreno evitar los viejos problemas. Las semánticas del contenido heredan los problemas que se les

presentaron a estas semánticas para el lenguaje natural. Ya hemos visto varios de ellos que, en última instancia, sólo son una repetición de los problemas que aquejan al significado/referencia de las expresiones lingüísticas: las disputas acerca de la determinación del significado, la del holismo-atomismo del significado, la del regreso de las interpretaciones para el caso de las teorías pictóricas, y demás.

Hasta aquí sólo he planteado el problema, tal como he mencionado mi intención no es elaborar ninguna semántica novedosa ni mucho menos una semántica para las representaciones mentales porque no creo que sea necesaria una semántica del contenido, al menos, no en el sentido filosófico tradicional y técnico del término "semántica". Pero sí diré algo, en el capítulo siguiente, acerca del marco conceptual a partir del cual entender la noción de representación mental.

Pasando al tema de la concepción estrecha de naturalización, además de lo que acabamos de ver acerca de la concepción referencial de la relación de representación, hay motivaciones ontológicas para sostenerla. En particular, se piensa que si no es posible dar condiciones naturales suficientes para que un estado físico tenga contenido, entonces los estados mentales no forman parte del mundo natural/causal y la consecuencia es el irrealismo acerca de lo mental. Pasajes como el siguiente plantean la preocupación:

Supongo que, más tarde o más temprano, los físicos completarán el catálogo que han estado compilando de las propiedades últimas e irreducibles de las cosas. Cuando lo hagan, cosas similares a *espín*, *encanto*, y *carga* aparecerán, quizá, en su lista. Pero *el ser acerca de*, seguramente, no aparecerá. Sencillamente, la intencionalidad no llega a esas profundidades. Es difícil ver, a la vista de estas consideraciones, cómo se puede ser realista acerca de la intencionalidad sin ser también, en mayor o menor medida, un reduccionista... la motivación más profunda para defender el irrealismo intencional se deriva ... de ciertas intuiciones ontológicas: que no hay lugar para las categorías intencionales en una imagen fisicalista del mundo; que lo intencional no puede ser *naturalizado*. (Fodor, 1987, p. 144)

El proyecto de naturalización del PNC consiste en lograr una teoría del contenido que articule en términos utilizados por las ciencias naturales condiciones

suficientes para que una representación sea acerca de algo en el mundo. O de otra forma, obtener una explicación de '*R representa S' es verdadero si C*, en donde la condición C se expresa con términos que forman parte de las ciencias naturales. La idea sería encontrar una condición C para cada estado mental-tipo, tal que si se obtendría C se instanciaría un estado mental-tipo. Se supone que si no es posible lograr esto, entonces, lo mental no forma parte del mundo natural, esto es, lo mental no existe.

Esta manera de plantear el problema tiene un doble defecto. Por un lado, si no es posible obtener condiciones suficientes especificables en términos de las ciencias naturales, no se sigue el irrealismo de lo mental. Porque si no es posible expresar esas condiciones en términos no-intencionales eso no dice nada acerca de la inexistencia de las condiciones que harían verdaderas expresiones como "PERRO tiene el contenido *perro*" (esto es, no implica la inexistencia de perros en el mundo). Y, si no es posible especificar condiciones (naturales o no) para que PERRO tenga el contenido *perro* esto no significaría que PERRO no se aplica a los perros (y por ende, que no sea verdadero que PERRO tiene el contenido *perro*) sólo significaría que hay ciertos conceptos (o términos) para los cuales no habría definiciones en términos de condiciones suficientes.<sup>6</sup>

Por otro lado, este planteo hace creer que la única manera de incluir lo mental en la imagen fiscalista del mundo es por medio de la búsqueda de condiciones suficientes naturales. Sin embargo, hay otra manera de concebir el proyecto de naturalización. Antes de pasar a otra concepción, hagamos un recorrido por las alternativas tradicionales al naturalismo del PNC.<sup>7</sup> A decir verdad, las alternativas tradicionales conciben el proyecto de naturalización del mismo modo estrecho que el PNC, es decir, todas son reductivas, y tienen tantas o más dificultades que el naturalismo (parcialmente) reductivo del PNC. Todas se enmarcan en la idea de que los estados mentales tienen esencias especificables en términos de condiciones necesarias y suficientes, y se distinguen entre sí porque algunas entienden la

---

<sup>6</sup> Véase Stich (1996, cap. 5) quien distingue tres tipos de naturalismos (basados en el análisis conceptual, las esencias y la relación de superveniencia) y muestra que del fracaso de cualquiera de ellos no se sigue el irrealismo de lo mental.

búsqueda de las esencias de una manera completamente *a priori*, otras pretenden ser compatibles con los hallazgos de la ciencia y otras aceptan sólo lo que la ciencia haya descubierto.

En primer lugar, está el proyecto tradicional de descubrir de manera *a priori* condiciones necesarias y suficientes para la aplicación de los términos mentalistas, apelando a las intuiciones de los usuarios de esos términos (Lewis 1972, Stalnaker 1984). El problema básico de este enfoque es que es sabido que el proyecto de análisis conceptual no dió resultado porque los términos no pueden definirse especificando condiciones necesarias y suficientes.

En segundo lugar, el proyecto de descubrir esas condiciones es en parte *a priori* y en parte empírico. Habría una etapa de reflexión *a priori* acerca del tipo de esencia de los estados mentales, por ejemplo esencias funcionales, y una etapa empírica que le corresponde a las ciencias y consiste en descubrir los ocupantes de esos roles funcionales (Dennett 1969, Field 1978). El problema básico aquí es que, dada la realizabilidad múltiple, no sería posible encontrar esencias neurofisiológicas, esto es, un estado neurofisiológico común a todas las criaturas que tengan instanciaciones de una creencia-tipo.

En tercer lugar, las esencias se descubren de manera totalmente empírica siendo a las ciencias y a la filosofía que teoriza sobre ellas a las que les corresponde su búsqueda, de modo que si los estados mentales no tienen ninguna de las propiedades atribuidas por el sentido común mucho peor para este último (Lycan 1988). Las objeciones a este tipo de proyecto suelen acudir a la idea de que si no hay restricciones conceptuales acerca de lo que se va a buscar, cualquier respuesta valdría, las esencias podrían ser tanto físico-químicas como psico-funcionales, y no habría razones para preferir unas a otras (Tye 1992). El problema con este enfoque no es esta objeción porque de hecho lo que se pretende con ella es que la búsqueda esté guiada por las intuiciones preteóricas de sentido común. La cuestión más bien radica en que no es necesaria la búsqueda de condiciones naturales para que la intencionalidad sea un fenómeno natural.

---

<sup>7</sup> Las alternativas se basan en la clasificación de Tye (1992) que no es igual a la de Stich (1996) aunque recoge los dos primeros tipos de naturalismos.

A continuación intentaré esbozar las líneas principales de una concepción naturalista alternativa. Las tesis básicas del PNA surgirán de la discusión de las tesis del PNC, desde las perspectivas de la tercera y cuarta estrategias.

## 2. El naturalismo amplio

A mi entender, y tal como se mencionó antes, lo que está en discusión es la peculiar concepción del naturalismo en la que se asienta el proyecto del naturalismo científicista (en particular las tesis (1) y (2)). Pero esta concepción parece ser, en parte, la consecuencia de las tesis acerca de la naturaleza de los estados intencionales (tesis (4)), y más precisamente, acerca del representacionalismo intencional (tesis (5)). En el §1.3 me he ocupado de las motivaciones semánticas y ontológicas que llevan a desarrollar una semántica naturalista reduccionista, aquí me concentro en las motivaciones metafísicas acerca de la concepción de los estados mentales (que, a decir verdad, le debe mucho a las motivaciones semánticas y ontológicas). Lo que intentaré mostrar es que la concepción metafísica de los estados intencionales es la que lleva a desarrollar no sólo una semántica del contenido sino una semántica naturalizada. La idea que subyace a este diagnóstico consiste en que si es posible concebir los estados intencionales de otra manera, entonces este tipo de naturalismo científicista no haría falta.

Si se parte de una postura antidualista, y se cree que hay estados mentales que están conectados con el mundo natural, se abren dos opciones acerca de cuáles son los elementos que intervienen en la conexión. Estas opciones dependen de qué concepción de los estados mentales se tenga. La primera opción es la de los defensores del PNC que creen que existen estados internos que poseen un contenido intencional, constituido por representaciones que poseen propiedades intencionales. Así, puesto que las propiedades intencionales de los estados mentales se asientan en las representaciones mentales, la intencionalidad es una cuestión de conectar las

representaciones con algo en el mundo. Si se piensa que la intencionalidad es una cuestión de vincular representaciones con algo del mundo, entonces surge la semántica del contenido mental cuya tarea es explicar cómo las representaciones internas adquieren contenido. Y puesto que se es antidualista, lo que hay que naturalizar es eso mental que está en el interior del individuo. Esto es así porque tener un contenido no es una propiedad física y la ontología del mundo es física, o la que establezca las ciencias naturales, de manera que hay que explicar esta propiedad en términos no-intencionales para que forme parte de la ontología del mundo.

Así, la tesis (4) del realismo intencional y la (5) del representacionalismo intencional llevan a la peculiar concepción del naturalismo especificada en las tesis (1)-(3). Estas tesis, en su conjunto, sostienen que hay algo así como representaciones intencionales, que constituyen el contenido de los estados intencionales, y que adquieren la propiedad de la intencionalidad por estar conectadas con lo que las ciencias naturales dicen que es el mundo natural (y, tal como se dijo en el capítulo anterior (1) y (2) implican (3)). De aquí el proyecto de una semántica naturalizada para los contenidos intencionales.

Quisiera hacer algunas observaciones con respecto a lo dicho. Si bien creo que la concepción de los estados mentales del PNC (la idea de que el contenido intencional está constituido por representaciones con propiedades intencionales) es condición suficiente y necesaria para el desarrollo de una semántica naturalizada, aquí me contento con exhibir algo mucho más débil, que consiste en mostrar la dicotomía entre representaciones intencionales - semántica naturalizada del contenido, por un lado y representaciones no-intencionales - no semántica naturalizada del contenido, por el otro. Esta dicotomía puede rastrearse en la literatura filosófica sobre el tema. Allí se puede apreciar que toda teoría que postula que el contenido está conformado por *representaciones con propiedades intencionales* desarrolla una semántica representacional de manera de dar cuenta de cómo esas representaciones adquieren su contenido a partir de cierta relación con el mundo.<sup>8</sup>

---

<sup>8</sup> Por supuesto esto se puede ver en los que defienden una u otra forma de la TRM y que han desarrollado o adhieren a alguna de las semánticas para la TRM: la semántica del rol conceptual, la

Se podría pensar que desarrollar una semántica no tiene por qué conllevar su naturalización, puesto que una cosa es conectar las representaciones con el mundo y otra es que esa conexión tenga que hacerse en términos no-semánticos. Sin embargo, no es el caso del PNC puesto que asume la tesis fisicalista. Es decir, si se cree que la última piedra de toque de lo mental radica en las representaciones mentales, no se querrá dejarlas desconectadas del mundo natural, puesto que hacerlo llevaría inevitablemente a un dualismo. Ahora bien, dada una semántica naturalizada, se podría pensar que no hace falta que esa naturalización sea científicista, en el sentido que los términos utilizados sean los de la ciencia natural. Y es cierto, porque se podría hacer uso de la terminología de una ciencia no-básica, pero las ciencias no-básicas (las ciencias humanas y sociales) tienen el mismo estatus que la psicología, por ende, no parece haber otra opción antidualista más que un naturalismo científicista para las representaciones mentales.<sup>9</sup>

De manera que, un recorrido posible que puede advertirse en el PNC es el que lleva desde la postulación de representaciones mentales intencionales a una semántica del contenido intencional, la que, a su vez, lleva al naturalismo, el que a su vez lleva a un naturalismo científicista. En cambio, y esta es la segunda opción, si se sostiene que, desde un punto de vista ontológico, un individuo posee estados intencionales (nadie niega que tenemos creencias, deseos, y similares), es decir, se es realista intencional, pero no se los concibe como una relación entre un individuo y representaciones mentales que son *portadoras de propiedades intencionales*, entonces no se hace necesaria la tarea de una semántica para el contenido ni mucho menos la de naturalizarlo.

Hasta aquí, he hecho el diagnóstico de por qué la concepción metafísica de los estados intencionales del PNC lleva a desarrollar una semántica naturalizada del contenido. Ahora, habría que decir algo, al menos, con respecto a dos

---

semántica informacional y la teleosemántica. Por otro lado, hay teorías de la intencionalidad que sostienen que hay representaciones, pero la intencionalidad es una propiedad del sistema cognitivo global (la persona), y en ese sentido, de los estados mentales personales, de manera que no se requiere de una semántica para las representaciones (es así como entiendo la postura de D. Dennett).

<sup>9</sup> Cf. Fodor (1994, p. 5) quien considera que el problema de la conexión entre las propiedades físicas y las no-físicas afecta a cualquier ciencia especial, esto es, a todas las ciencias menos la física que, en sentido estricto, es la ciencia básica.



cuestiones. La primera es mostrar por qué esta concepción es infructuosa, y la segunda, consiste en delinear una opción para dar cuenta del fenómeno de la intencionalidad dentro de un marco naturalista. La concepción de la naturaleza de los estados intencionales del PNC es infructuosa porque al asimilar dos tipos de fenómenos, el representacional y el intencional, deja sin explicar cuestiones fundamentales de cada uno. De manera resumida, la complejidad del fenómeno de la intencionalidad no gana en comprensión postulando representaciones intencionales, y la postulación de representaciones intencionales no aporta claridad sino que oscurece casi por completo el rol explicativo de las representaciones en las teorías cognitivas. En lo que sigue me ocuparé de por qué el PNC no explica lo propio del fenómeno intencional, y en el desarrollo de esto esbozaré una opción naturalista para abordar este fenómeno (opción que se completará de delinear al discutir el resto de las tesis del PNC). Si bien iré adelantando algunas ideas sobre por qué las teorías psicológicas no requieren de representaciones intencionales, todo el capítulo siguiente estará dedicado a esta cuestión.

Las semánticas del contenido simplifican tanto el fenómeno de la intencionalidad que realmente no queda claro qué es lo que están explicando, o al menos, si explican algo, eso que explican parece ser un cambio de tema. La simplificación se debe, en parte, tal como hemos visto, a la concepción de los estados intencionales y el consecuente proyecto de naturalización. Revisemos más detenidamente el proyecto para dejar en claro por qué se simplifica el fenómeno de la intencionalidad. Según el PNC, las representaciones tienen la propiedad de ser intencionales. La naturalización de esta propiedad relacional abarca una doble tarea simultánea para el logro de un único objetivo: explicar cómo lo mental se relaciona con lo físico. La tarea es doble porque obtener la explicación en cuestión da lugar a problemas específicos que se relacionan por un lado, con el problema mente-cuerpo y por el otro, con el problema de la relación mente-mundo externo. Veamos primero esto último.

Por un lado, se requiere explicar cómo ciertos estados físicos se vuelven intencionales o, para ser más exactos, representacionales.<sup>10</sup> La intencionalidad se la concibe como la relación representacional que conecta propiedades mentales con propiedades físicas del mundo, de manera que las primeras sean *acerca de* las segundas.<sup>11</sup> Según este proyecto de naturalización, una propiedad es natural si y solo si está relacionada de manera legítima con propiedades físicas del mundo. Llamemos a esa relación legítima R1. En el caso de los estados mentales (pero lo mismo valdría para dar cuenta de, lo que desde un punto de vista intuitivo, serían propiedades que van más allá de las físicas como ser un caballo, un planeta, una sociedad, y demás), R1 es legítima si está formulada en términos (al menos, parcialmente) reduccionistas que apelen sólo a los que las ciencias naturales admiten. De manera que en R1 intervienen como *relata* las propiedades mentales y físicas. Estas últimas, dadas ciertas condiciones físicas suficientes, dan lugar a las primeras y las primeras refieren a las últimas. R1 se la concibe como la relación técnica de referencia, de la que hemos hablado en el §1.3.

Por otro lado, se requiere explicar cómo los estados intencionales se relacionan con los estados físicos del cerebro, esto es el problema mente-cuerpo. Lo que se busca aquí es una relación R2 que permita incluir a los estados intencionales en el mundo físico, sobre todo se busca legitimar el poder causal de los contenidos de los estados intencionales. El PNC es fisicalista y apuesta a un proyecto reductivo del *contenido* de los estados intencionales, no de los estados mismos. Es por esto

---

<sup>10</sup> Dicho de esta manera parece una cuestión temporal y no lógica/conceptual, como si al nacer sólo tuviéramos estados físicos (y, para algunos, las formas o sintaxis de un sistema representacional primitivo innato como el mentalés), y durante la ontogénesis adquiriéramos mentalidad. A decir verdad, el PNC no es explícito al respecto. Es presumible que la TDA (dada la postura de Fodor con respecto al mentalés) ubique en el aprendizaje de la lengua nativa la ocasión para el desarrollo de la mentalidad, pero no queda claro cómo sería esto en los sistemas cognitivos no-lingüísticos.

<sup>11</sup> En lo que sigue voy a plantear la discusión en términos de propiedades y relaciones causales. Pero, lo que se diga vale lo mismo para la teleosemántica. Por un lado, las representaciones articuladas refieren biunívocamente a propiedades del mundo. Por otra parte, se suele considerar a la teleosemántica una teoría no sólo histórica sino causal. De hecho, lo que permite que una representación se asocie históricamente con una condición del entorno es un proceso causal. De todas las conexiones causales que podrían constituir el contenido de una representación, la teleosemántica selecciona la adecuada en virtud del rol adaptacional (pero, lo que interesa para fijar el contenido no son las relaciones causales actuales sino aquella que fue seleccionada en la historia del ítem intencional). Cf. Quesada (1995), Cummins (1997).

que excluye entender a R2 como paralelismo, armonía preestablecida, interacción causal (todas relaciones dualistas) e identidad de tipos (relación reductiva entre estados mentales-tipo). Las relaciones que quedan, las que se han ofrecido en la literatura sobre el tema, son las que habitualmente adoptan los fisicalismos no-reduccionistas, a saber y principalmente, las de realización (o implementación) y superveniencia.<sup>12</sup>

La relación entre R1 y R2 es la siguiente: dado que la intencionalidad es una propiedad relacional de las representaciones, ésta *superviene* sobre propiedades físicas relacionales de (o extrínsecas a) la persona. Así, el contenido intencional superviene sobre propiedades físicas de la persona que se definen en función de las relaciones entre la persona y el mundo (R2), y los estados físicos son acerca de cosas en el mundo porque esas relaciones permiten que éstos *refieran* a cosas del mundo (R1).<sup>13</sup> Uno de los problemas fundamentales de esta explicación es que supone que sólo R2 puede dar cuenta de la conexión con los estados físicos de la persona y sólo R1 puede dar cuenta de la conexión con el mundo. En cambio, creo que sería más fructífero pensar que hay tanto un pluralismo de relaciones mente-cuerpo como un pluralismo de relaciones mente-mundo, esto es, un pluralismo metafísico y un pluralismo semántico. Y la razón para esto es que los *relata* de R1 y R2 no son homogéneos. Diré muy poco sobre R2 porque lo que interesa aquí es el aspecto de la intencionalidad que es el de 'ser acerca de' y me concentraré en argumentar que concebir a R1 como una relación de referencia simplifica tanto el problema de la relación mente-mundo que no se entiende en absoluto cómo una persona se las arregla para que, a través de su pensamiento, conciba o interprete cómo es el mundo de manera de poder hablar acerca de él y realizar además otros tipos de acciones.

---

<sup>12</sup> Para la relación de realización ver nota 23 del cap. 1 y §2.2 del capítulo 2. La idea principal de la relación de superveniencia es que "los estados del tipo *X* supervienen en estados del tipo *Y*, si y sólo si no hay una diferencia entre los estados *X* sin que haya una diferencia correspondiente entre los estados *Y*" (Fodor 1987, p. 55).

<sup>13</sup> Si R2 es la relación de realización o implementación, el contenido intencional que está relacionado con el mundo por R1, se realiza o implementa en los estados físicos de la persona por R2.

Con respecto a la homogeneidad de los *relata* de R1 y R2, el PNC asume que las entidades físicas comparten una misma naturaleza física, esto es, que son ontológicamente homogéneas. No importa que se esté hablando de estados físicos de una persona, de objetos físicos en el mundo, o lo que sea físico, estos tipos de entidades comparten cierta naturaleza. En general, se suele afirmar que la naturaleza de las entidades físicas es la que la ciencia física especifica. Además, se asume que hay entidades que son puramente o meramente físicas como distintas de otras que además tienen otro tipo de propiedades no-físicas como por ejemplo, las propiedades mentales. Con respecto a lo mental, se suele asumir lo mismo. Todas las entidades mentales, más allá de las distinciones que se puedan hacer (entre estados, procesos, etc.), comparten, en un sentido relevante tanto para el problema mente-cuerpo como para el problema mente-mundo, una misma naturaleza. Por ejemplo, hemos visto que para el funcionalismo, lo que tienen en común las entidades mentales es su esencia funcional. Además, se asume que los estados intencionales están constituidos también de una manera homogénea por representaciones intencionales.<sup>14</sup>

Sin embargo, no está claro qué es una propiedad física, y además bajo cualquier noción intuitiva de lo físico, tampoco está claro que haya objetos que sólo tengan propiedades físicas a diferencia de objetos que tienen además otros tipos de propiedades. Creer que se puede distinguir “entre las cosas físicas que también son mentales y aquellas que son *meramente* físicas” (Lycan 2003, p. 16) parece ser una tarea tan infructuosa como querer distinguir entre las cosas físicas que también tienen propiedades eléctricas, químicas, biológicas, ópticas, sociales, funcionales y aquellas que son meramente físicas. Todo objeto que intuitivamente se podría considerar como meramente físico, desde una silla a un insecto, tiene propiedades no-físicas, y en este sentido, no debe existir nada que sea meramente físico.<sup>15</sup>

---

<sup>14</sup> Ver Chomsky (1995b, 2003) para las dificultades de definir lo físico y van Gelder (1998) para problemas con la concepción de lo mental.

<sup>15</sup> El contexto de la cita es el siguiente:

En particular, las cosas físicas que representan el mundo tienen una gama completa de propiedades interesantes que las cosas meramente físicas no tienen, tales como valores de verdad y relaciones de implicación. Los organismos que tienen mentes (no obstante, físicas)

Por otro lado, al no tener una formulación adecuada de qué son las propiedades no-físicas no queda claro por qué decir que aquellas cosas que tienen propiedades mentales, eléctricas, sociales o, funcionales no son 'meramente' físicas. El problema aparece también del lado de lo mental. No está claro la naturaleza de los contenidos intencionales y su expresión en términos semánticos/intencionales. Afirmaciones como la siguiente resultan sospechosas cuando todo un proyecto de naturalización se basa en la distinción entre propiedades (o predicados) intencionales y físicos: "dado que no tenemos ninguna manera general y satisfactoria de decir cuáles expresiones *son* semánticas(/intencionales), queda a la intuición determinar cuándo una formulación de C [la condición natural suficiente] cumple esta condición [la de estar formulada en términos no-semánticos y no-intencionales]" (Fodor 1990c, n. 1, p. 48).

Todos estos problemas con respecto a la falta de una idea clara de qué es lo físico y lo mental, ya de por sí llevaría a pensar que el tipo de naturalismo metafísico sustancial del PNC es informulable. Hemos visto en el §1.3 que las alternativas naturalistas al PNC también tienen problemas. Pero el problema fundamental es que se basan en una distinción entre propiedades (entidades, hechos, predicados, etc.) físicas y mentales difícilmente formulable. De manera que todo intento de naturalización basado en la idea de que 'no hay propiedades mentales además de las propiedades (entidades, hechos, etc.) físicas' está destinado al fracaso.

Stich (1996) y Tye (1992) llegan a esta conclusión por otros caminos que el abordado aquí (aunque ambos autores lo hacen por razones bastantes similares) y

---

son capaces de agencia, que les provee aún de otra amplia gama de rasgos interesantes que no están disponibles para lo meramente físico, etc. (Lycan 2003, p. 16)

El problema de la distinción no es sólo el mencionado en el cuerpo del texto, sino el de cómo entender 'lo mental'. Esta cita asume que las representaciones son intencionales, que tener representaciones constituye el ámbito de lo mental y que lo mental define tener una mente. Sin embargo, hay objetos estudiados por la psicología cognitiva animal, a los que se les atribuye representaciones, por ejemplo las que se atribuyen al sistema computacional que subyace a la navegación de los insectos o al canto de los pájaros. ¿Esto quiere decir que tienen una mente? ¿que sus representaciones son intencionales? ¿que son agentes? El problema radica en colocar a las representaciones, lo intencional, y la agencia en el mismo paquete de lo mental. Como quedará claro a lo largo de este capítulo y del siguiente, tener representaciones no es suficiente para la intencionalidad y la agencia (se requiere de prácticas lingüísticas).

proponen una alternativa naturalista.<sup>16</sup> Stich afirma, con un criterio pragmatista, que la naturalización de lo intencional no pasa por una respuesta metafísica sino por tener una teoría científica exitosa que invoque propiedades intencionales. Tye, reformulando las nociones de "mental" y "mundo natural", afirma, con un criterio teórico-metafísico, que la naturalización de lo mental pasa por considerar a la psicología como una ciencia física. En realidad, ambas propuestas, ya sea con criterios pragmáticos o metafísicos, siguen apegadas al cientificismo del PNC. Es cierto que en la propuesta de Stich no se apela a una reducción de las propiedades intencionales a las de las ciencias naturales, pero la idea de que sólo es natural aquello que participa de una teoría científica tiene un resabio cientificista, además de los problemas acarreados por el criterio pragmatista: ¿cuáles son las disciplinas consideradas ciencias? ¿qué sucede si ninguna ciencia invoca propiedades intencionales, esto querría decir que no son naturales?

Creo que el camino adecuado para delinear las bases de un naturalismo metafísico sustancial es, al igual que Tye, modificar la concepción de los estados intencionales y la de mundo natural. Pero, a diferencia de Tye, considero que la tarea se puede realizar sin que conlleve intervención alguna de las teorías científicas.<sup>17</sup> En lo que sigue, intentaré mostrar que R1 no es la relación adecuada para dar cuenta de la conexión entre la mente y el mundo. En parte, como se dijo antes, esto es así por la concepción homogénea de los estados mentales como estados intencionales compuestos por representaciones con propiedades intencionales y además por la concepción que se tiene del otro miembro de la relación, esto es, el mundo.

---

<sup>16</sup> También Horgan (1992) expresa sus dudas acerca de la posibilidad de especificar condiciones suficientes no-semánticas y no-intencionales para los estados intencionales, inclinándose a una salida alternativa naturalista basada en una explicación natural de la relación de superveniencia de lo intencional sobre lo físico.

<sup>17</sup> En particular porque no considero, tal como se verá en el capítulo siguiente, que en la psicología cognitiva intervengan nociones (propiedades, predicados, etc.) intencionales. Eso no significa ni que la psicología no pueda aportar sus explicaciones acerca de los mecanismos subpersonales para echar luz sobre aspectos del fenómeno de la intencionalidad ni que éste no pueda ser abordado por un conjunto de disciplinas científicas. Sólo significa que la naturalización no depende de que la psicología sea intencional ni de sus relaciones con las ciencias naturales.

La relación R1 se establece entre instancias de representaciones mentales e instancias de propiedades en el mundo.<sup>18</sup> Un primer problema es que el contenido de nuestras creencias no parece corresponderse con la individuación de una cierta propiedad en el mundo. Un segundo problema es que el contenido de nuestras creencias no se agota identificando una cierta propiedad en el mundo. Con respecto a la primera cuestión, R1 supone que cuando pensamos AGUA, esta instancia de representación está conectada con la propiedad en el mundo H<sub>2</sub>O. Este es un ejemplo paradigmático porque “agua” es un término de clase natural para el que se supone que hay esencias. Es cierto que una de las propiedades del agua es ser H<sub>2</sub>O pero también lo es ser un líquido incoloro, inodoro, y demás, o la de ser un líquido que sale de las canillas, o el que compone los mares, ríos, y similares. Si pensamos en el agua del Río de la Plata difícilmente esté compuesto de H<sub>2</sub>O o si se tira un saquito de té en un tanque de agua y se distribuye a los hogares, seguiríamos pensando que lo que estamos tomando es agua aunque fuera un tanto extraña.

Por más que se considere que la propiedad esencial de “agua” es H<sub>2</sub>O, no siempre es en lo que pensamos cuando pensamos en AGUA. De manera que no parece haber correspondencia entre lo que pensamos y los referentes individuados por la relación R1. Por otro lado, hay términos como “libro” para los que no parece haber propiedades esenciales, ¿con qué propiedad esencial tendríamos que estar conectados para que adquiriera significado? Cuando pensamos LIBRO asociamos una serie de propiedades, algunas de las cuales parecen incluso incompatibles a primera vista. Un libro puede ser concreto (en tanto ejemplar) pero también abstracto (en cuanto tipo) como cuando decimos que estamos leyendo el mismo libro que está leyendo otra persona.<sup>19</sup>

Con respecto a la segunda cuestión, nuestras clasificaciones de objetos, sustancias, acciones, y demás dependen de nuestros intereses y propósitos. Así, podemos clasificar por la constitución material, por la función, por su rol

---

<sup>18</sup> O entre representaciones mentales articuladas y estados de cosas, en donde cada elemento del primero refiere unívocamente a un elemento del segundo. Como se mencionó en la nota 11, la relación de referencia está expresada en términos de propiedades pero los argumentos son extensibles a la teoría pictórica.

<sup>19</sup> Los ejemplos de “agua” y “libro” son los que suele utilizar Chomsky (por ejemplo, 1998, pp: 51 y ss.).

institucional, etc. Cada vez que pensamos en algo en particular, no siempre estamos interesados en la propiedad esencial de la cosa (si es que hay de eso). Cuando pensamos AGUA en un contexto particular hay una serie de creencias asociadas que responden a nuestros intereses y propósitos que parecen participar en la individuación de ese contenido. Decidir que la propiedad referida para ciertas sustancias es la constitución química no deja de ser un criterio clasificatorio en función de ciertos intereses, muy probablemente, científicos.

Ambos problemas llevan a la idea de que si se reduce el fenómeno de la intencionalidad a R1 se están dejando de lado elementos de la relación que parecen participar ineludiblemente en el fenómeno de la intencionalidad: ésta no es una relación de dos elementos (contenido-aspecto del mundo) sino de cuatro elementos: es una persona quien piensa algo por medio de una expresión bajo ciertas circunstancias (persona, mundo, expresión, contexto). La idea de una persona que refiere a aspectos del mundo usando expresiones bajo cierto contexto es ya la puesta en uso del lenguaje y con ello el ámbito de la comunicación. En la comunicación atribuimos y se nos atribuyen estados intencionales con contenido y lo hacemos desde cierta perspectiva que adoptamos, focalizando nuestros intereses en determinados aspectos.

Lo que pensamos, creemos, y similares no es una cuestión de qué representaciones tenemos instanciadas en nuestros cerebros y con qué cosas del mundo físico se conectan, sino que la cuestión de la intencionalidad radica en que son las personas, que están inmersas en el mundo natural, las que a través de sus intereses, puntos de vistas y, en general de las prácticas comunicativas significan (/refieren) las expresiones lingüísticas que manifiestan creencias, deseos, y similares. Los estados personales intencionales no son relaciones con representaciones intencionales que refieren a aspectos del mundo sino que son la interiorización de expresiones lingüísticas. El pensamiento es una actividad lingüística que realizan las personas en función de cierta perspectiva, intereses, etc. Son las prácticas comunicativas, la puesta en uso del lenguaje, las que nos permiten pensar y hablar acerca del mundo desde cierta perspectiva. Estas perspectivas de



las personas se relacionan con las creencias asociadas a los términos lingüísticos.<sup>20</sup> Así hay elementos epistémicos que intervienen en la individuación de lo que pensamos.

Estos elementos epistémicos, que son la perspectiva de la persona, se relacionan con cómo el agente interpreta el mundo. En nuestras prácticas comunicativas de atribución de estados intencionales tomamos en cuenta las perspectivas de las personas. De hecho, atribuimos creencias y deseos en función de lo que sabemos acerca de las personas, tanto en términos generales e intuitivos de qué es ser una persona, como en términos particulares de lo que sabemos de esa persona. Por ejemplo, si el contexto de atribución es el científico y la persona a la que le vamos atribuir estados intencionales para hacer inteligible su comportamiento es un químico, seguramente tendremos en cuenta su perspectiva y probablemente el contenido de sus pensamientos acerca de agua sea  $H_2O$ . Mientras que al resto de las personas les atribuiremos contenidos en función del contexto comunicativo, esto es, si estamos hablando acerca del Río de la Plata o de los tanques de agua y demás, y de lo que sepamos de ellas.

Esto no significa que tengamos que tener ninguna teoría interna acerca de la atribución intencional, ni que estemos constreñidos por principios de racionalidad en la atribución intencional (normatividad de lo mental), ni que tengan que haber normas que gobiernen los usos lingüísticos (normatividad lingüística). Predecimos más o menos adecuadamente el comportamiento de los demás simplemente porque las personas expresan sus creencias y deseos, conocemos sus perspectivas, proyectamos patrones de conducta observados en el pasado y suponemos ciertas regularidades lingüísticas de manera de poder interpretar sus expresiones.<sup>21</sup>

---

<sup>20</sup> La noción de perspectiva de las personas es en el sentido de Bilgrami (1992).

<sup>21</sup> Las regularidades lingüísticas no suponen normas sino expectativas compartidas que son algo así como guías pragmáticas en el sentido davidsoniano o de Bilgrami (1992). Davidson (1984, 1986) niega que el significado esté gobernado por normas, reglas o convenciones. La única norma, si se la quiere llamar así, está dada por la intención del hablante de ser interpretado de cierta manera (su deseo de comunicarse), así "el hablante tiene que 'proseguir' más o menos como los otros esperan o al menos están equipados para interpretar" (Davidson 1992, p. 115). En este sentido, se trata de una cuestión práctica. Simplemente y en general, queremos ser entendidos. Se trata de lo que el hablante tendría que hacer si quiere ser entendido en contraposición a sostener que el significado tiene un aspecto esencial de corrección o incorrección (normatividad del significado). La noción de

Inducciones, perspectivas de las personas y principios pragmáticos comunicativos es todo lo que necesitamos para la atribución intencional.

En este sentido, nuestra conexión con el mundo natural no pasa por vincular representaciones con aspectos del mundo sino que la conexión se da en el uso del lenguaje, en el ámbito de la comunicación. De manera que, nuestros estados personales intencionales expresados en las expresiones lingüísticas se relacionan con el mundo a través de nuestras prácticas lingüísticas comunicativas. Estas prácticas comunicativas están ancladas en "formas de vida" (en el sentido de Wittgenstein 1953, #241) que se edifican sobre la base de condiciones naturales: un "comportamiento común de la humanidad" (Wittgenstein 1953, #206), "un sistema preexistente de percepciones" (Pears 1995) y "regularidades en la naturaleza" (Hacker 1972, p. 297).

Habíamos dicho, al principio de esta segunda parte del capítulo, y hemos visto, que la idea de conectar representaciones con el mundo para dar cuenta de la intencionalidad de los estados de actitud proposicional llevaba a la simplificación del fenómeno de la intencionalidad porque dejaba a un lado elementos indispensables para comprender cómo podemos pensar acerca del mundo. Pero, también habíamos dicho que la postulación de representaciones intencionales no ayudaba a entender el rol explicativo de las representaciones en las teorías cognitivas. De esto me ocuparé en el capítulo siguiente, pero a modo de adelanto y para tener una idea global sólo diré que el PNC no recoge el rol explicativo de las representaciones en (al menos, algunas de las teorías de) la psicología cognitiva, dado que la noción de representación que se postula no es intencional.

Las semánticas del contenido o psicosemánticas se ofrecieron para complementar a la TRM, que se postuló como aquella teoría que subyacía a las teorías psicológicas. Sin embargo, las representaciones tienen un rol explicativo en las teorías de la psicología cognitiva que no parece depender de una psicosemántica. Es más, la psicosemántica no parece echar ninguna luz para dar cuenta de los problemas para los cuales se postularon representaciones y, en

---

corrección, en esta perspectiva, es secundaria a la del deseo de comunicarse (Cf. Bilgrami, 1992, cap. 3).

cambio, agregan problemas, para los cuales las teorías psicológicas no fueron pensadas para resolver.<sup>22</sup> El punto es que las teorías psicológicas más desarrolladas actualmente se mueven en el ámbito subpersonal, son intentos de dar cuenta de cómo funcionan los mecanismos subpersonales (como el de la visión o lenguaje), y la noción de representación está relacionada con la idea de información interna a los sistemas subpersonales. Esa información (en un sentido que se explicitará en el capítulo siguiente) no parece ser acerca del mundo.

De manera que, por un lado, (si se acuerda conmigo en los argumentos del capítulo siguiente) las representaciones que intervienen en los estados y procesos subpersonales no son las que hay que conectar con el mundo, y por otro lado, las que intervienen en los estados personales están conectadas no porque tengan propiedades intencionales (o sean las portadoras de la intencionalidad) sino porque son las personas las que están conectadas con el mundo a través de las prácticas comunicativas. De manera que, para el PNA, el lugar del fenómeno de la intencionalidad es en el ámbito de la comunicación y la intencionalidad es un fenómeno natural porque está anclado en condiciones de vida naturales. Por supuesto que ambas afirmaciones requieren de ciertas especificaciones.

Con respecto a la primera, la manera en que nuestras expresiones adquieren significado, la naturaleza del vínculo lenguaje-mundo es algo que queda por explorar porque hay que tener bien en claro los elementos de la relación hablante-expresión-contexto-mundo: qué aspectos de los hablantes intervienen, la naturaleza de las distintas expresiones lingüísticas, los criterios de distinción de contextos, qué aspectos del mundo se toman en cuenta, y demás. Aquí se podría optar por cualquier semántica o, mejor dicho, pragmática, que se considere satisfactoria para dar cuenta de esta relación. Sin embargo, lo que se podría sugerir es que quizá así como parece más fructífero pensar que no hay un solo tipo de conexión entre los estados mentales y los estados físicos, podría ser que no hubiera

---

<sup>22</sup> Ver §1 del capítulo 2 para lo primero. Los problemas que parecen interesarles a los científicos se relacionan con cuáles son las propiedades de los sistemas computacionales encargados de las capacidades cognitivas, cómo se relacionan entre sí, etc. La introducción de nociones intencionales no ayuda en nada sino que más bien plantea toda una gama de problemas, algunos de los cuales hemos estado viendo hasta aquí y otros los veremos en el capítulo siguiente.

un solo de tipo de conexión con el mundo sino distintos dependiendo de los diferentes aspectos de los elementos que intervienen en la relación. En este sentido, quizá la opción más razonable, tal como se mencionó antes, sea tanto la de un pluralismo metafísico como un pluralismo semántico/pragmático. Con respecto a la segunda, una vez que el mundo natural no es el estrecho de las ciencias naturales, podemos considerar al aspecto mental de las personas como un aspecto tan natural como los aspectos químicos, eléctricos, y demás.

Se podría resumir lo dicho hasta aquí en las siguientes tesis:

1- *Naturalismo*: los pensamientos, las creencias y similares de las personas forman parte del mundo natural porque están anclados en prácticas comunicativas que a su vez, se anclan en “formas de vida”, las que a su vez están ancladas en sistemas senso-perceptivos, y ciertas regularidades físicas y comportamentales.

3- *Realismo intencional*: existen estados internos en las personas que son aquellos a los que se alude en la atribución intencional por medio de oraciones de AP. La naturaleza de estos estados refleja la naturaleza lingüística de las oraciones de AP.<sup>23</sup>

Voy a ocuparme ahora de la discusión del resto de las tesis acerca de la naturaleza de los estados intencionales (tesis (6), (7) y (8)). Aunque no haya conexiones directas entre las representaciones internas y algo específico en el mundo que fije su contenido, la tesis (6) del externalismo puede seguir en pie. Lo que pensamos, expresado en nuestras emisiones lingüísticas (manifiestas o no), está determinado por nuestras relaciones con el mundo, éste contiene tanto el mundo de los objetos físicos como el de nuestras prácticas comunicativas que están ancladas en nuestros sistemas senso-perceptivos y en ciertas regularidades. Tal como se mencionó antes, cuáles sean los procesos por los cuales nuestras expresiones (y con ello los estados mentales personales intencionales) adquieren significado es una cuestión empírica a explorar que le compete al conjunto de las disciplinas encargadas de dar

---

<sup>23</sup> Los números de las tesis del PNA responden al orden de su presentación en conjunto al final del capítulo. Para evitar confusiones los números de las tesis del PNC, a diferencia de las del PNA, siempre aparecen entre paréntesis.

cuenta de los procesos de comunicación. De manera que la tesis externalista se mantiene:

4- *Externalismo*: el contenido de los estados intencionales de una persona está determinado por sus relaciones con el mundo.

Con respecto a la tesis (7), para el PNC es fundamental que el contenido esté determinado unívocamente dado que R<sub>1</sub> es una conexión entre representaciones y aspectos del mundo, o en otras palabras, hay una correspondencia biunívoca entre elementos semánticos y ontológicos. Si no se puede establecer objetivamente esta correspondencia, esto es, si no hay un hecho objetivo en el mundo acerca del contenido de una representación, no habría condiciones naturales para que surja y todo el proyecto colapsaría. El ejemplo que ilustra la indeterminación del contenido es el que hemos visto en el §1.2 en relación al contenido de una rana que saca su lengua para atrapar moscas o puntos negros en movimiento. Es a este problema de la disyunción de los contenidos al que se alude cuando en la literatura sobre el tema se habla de indeterminación.<sup>24</sup>

Si el contenido de la representación de la rana depende de nuestra interpretación, entonces no está determinado objetivamente y es en este sentido que no habría un hecho decisivo en el mundo acerca de cuál es su contenido. El problema no trata acerca de la posibilidad de adscribir cualquier contenido sino acerca de establecer entre aquellos aspectos del mundo con los que una representación mantiene relaciones causales, aquel que constituye el contenido de esa representación. Esto no se soluciona argumentando que, en principio, estaría ontológicamente determinado aquel aspecto del mundo que juega el papel causal relevante y que sólo sería una cuestión epistemológica, por cierto compleja, averiguar cuál es. El punto es que una semántica del contenido debería proveer las herramientas necesarias para establecer cuál es el contenido, de otro modo, la teoría

---

<sup>24</sup> En la tradición filosófica hay, al menos, tres tesis clásicas de la indeterminación que surgen de los escritos de Quine: la indeterminación de la traducción, la inescrutabilidad de la referencia y la sobredeterminación de las teorías científicas (Cf. Quine 1990). El sentido de indeterminación que se discute aquí está relacionado, en algún sentido, con la segunda tesis.

no es satisfactoria, no cumple su objetivo, y por consiguiente, no habría manera de dar cuenta de cómo surge el contenido.

Si no se piensa la cuestión de esta manera, no se entiende por qué los defensores de las semánticas del contenido han intentado resolver el problema de la disyunción al punto que lo han transformado en el criterio que permite decidir si una semántica es satisfactoria (vuelvo a remitir al §1.2). Tal es la dificultad que según la teleosemántica el contenido de la rana es *mosca* y según la TDA es *punto negro en movimiento*. Esto es interesante porque muestra otro tipo de indeterminación, ambas semánticas son compatibles con la evidencia disponible pero incompatibles entre sí porque asignan distintos contenidos a las representaciones de un organismo.<sup>25</sup> Mientras que el problema de la disyunción sigue siendo un problema para el PNC, no lo es para el PNA.<sup>26</sup>

Es sabido que el problema de la indeterminación para el lenguaje surge básicamente cuando se sostiene que la relación lenguaje-mundo consiste en R1. Esto es, cuando se sostiene que los términos y predicados del lenguaje refieren a entidades (propiedades, y demás) particulares de la ontología del mundo.<sup>27</sup> Quine (1960) muestra que hay casos en los que no hay un hecho objetivo que permita decidir cuál es el significado de un término. Puesto que en el PNA la relación entre lenguaje-mundo no consiste en R1, no deberían por qué surgir los problemas de indeterminación. En el PNA, no se trata de que haya una propiedad en el mundo que sea "el" contenido de una representación con independencia de los intereses y los puntos de vista de los hablantes sino que el contenido se constituye de manera más compleja no sólo teniendo en cuenta el aspecto mencionado sino el resto de los elementos que intervienen en la relación hablante- contexto- expresión- mundo.

Dado que en esta relación participa la perspectiva del hablante, no hace falta que haya hechos objetivos *decisivos* que fijen el significado de una vez y para siempre

---

<sup>25</sup> Esto sería una forma de la tesis de la indeterminación de la traducción: "[las] afirmaciones de dos manuales podrían ser verdaderas y sin embargo, las relaciones de traducción de ambas podrían no ser usables alternativamente de una oración a otra, sin dar lugar a secuencias incoherentes" (Quine 1990, p. 5).

<sup>26</sup> El problema de la disyunción tiene consecuencias para el establecimiento de leyes psicológicas. Si resulta que las propiedades intencionales son propiedades disyuntivas abiertas, éstas no podrían figurar en leyes intencionales porque no son proyectables.

<sup>27</sup> Cf. Quine (1990), Davidson (2001) y, Hornstein (1991).

para todos los contextos de emisión. Sólo hace falta que haya un mundo o hechos o propiedades o lo que se considere del mundo que participa en la relación lenguaje-mundo. Cuando atribuimos contenidos, tomamos en cuenta la concepción del mundo de las otras personas de manera que podemos establecer, la mayoría de las veces, con bastante precisión cuáles son los contenidos de sus pensamientos (y en este sentido, no son los referentes en sí mismos los que individúan el contenido). Esto no quita que hay ocasiones en las que puedan coexistir interpretaciones alternativas sin que podamos decidirnos por una u otra, no obstante se puede tolerar perfectamente un poco de indeterminación dado el hecho efectivo de la comunicación.

Considerando que desde esta perspectiva, los contenidos no dependen de que haya un hecho decisivo en el mundo que determine acerca de qué son, la tesis de la determinación del contenido no es necesaria. De manera que si determinación quiere decir que sólo los hechos decisivos en el mundo fijan el contenido, el PNA sostiene que no son los hechos decisivos en el mundo los que fijan el contenido, sino que el mundo (o lo que sea de él) participa junto con otros elementos de la fijación del contenido, de modo que si no hay hechos decisivos, esto no constituye ningún problema para el PNA. A falta de una denominación mejor, llamaré tesis de la indeterminación a lo que sigue:

5- *Indeterminación del contenido*: no se requieren hechos decisivos con respecto a cuál es el contenido de un estado mental de una persona.

Con respecto a la tesis (8), mientras que en el PNC las creencias asociadas, con cierta creencia, no participan de la determinación ni individuación de los contenidos por las razones vistas en el capítulo anterior, el PNA sostiene que en la atribución intencional es ineludible la apelación a la perspectiva de la persona para individuar los contenidos de sus pensamientos. Las creencias y concepciones de la persona acerca del mundo median sus relaciones causales y permiten (la mayoría de las veces) individuar de entre aquellas las que constituyen el contenido. Así, Rosana y Claris cuando conciben el agua desde la perspectiva de un líquido incoloro, inodoro, que sirve para calmar la sed, y demás, comparten las mismas creencias acerca del agua.

Si Rosana es química y Claris no lo es, el acuerdo o no entre sus creencias dependerá de los contextos de emisión. Si el contexto es científico, seguramente tendrán distintas perspectivas porque Claris no sabe que el agua es H<sub>2</sub>O (esto no significa que H<sub>2</sub>O sea la propiedad determinante de la referencia de "agua" porque, tal como vimos, la relación de referencia es más complicada que lo que R1 hace suponer. La propiedad del agua de ser H<sub>2</sub>O es una de entre las muchas propiedades del agua). Pero si el contexto es el cotidiano, ambas compartirán perspectivas. Las creencias que participan efectivamente en una situación en particular no tienen por qué ser siempre las mismas, lo único que se requiere para la comunicación es un solapamiento considerable de las distintas perspectivas y eso ocurre porque compartimos un mismo mundo natural (de regularidades físicas, comportamentales y lingüísticas).

De aquí se sigue que hay elementos que intervienen en la individuación del contenido que dependen de las creencias asociadas que recogen los intereses, puntos de vista y la concepción del mundo de las personas que las poseen, por ende la tesis (8) del atomismo del contenido no se sostiene. La tesis más adecuada sería la siguiente:

6- *Molecularismo del contenido*: la identidad del contenido de un estado intencional depende de los estados intencionales asociados en contextos particulares.

Con respecto a la idea de que la PSC es el punto de partida para una psicología científica (tesis (9)) y la idea de que una psicología científica requiere una semántica del contenido naturalista y científicista (tesis (10)), no creo que esta sea una manera fructífera de concebir la tarea y los objetivos de la psicología cognitiva. Estas tesis se basan en la idea de que la psicología tienen como objetivo explicar la conducta intencional por medio de explicaciones intencionales. En el capítulo 1 he dado algunos argumentos de por qué no considero que esa sea la tarea de la psicología y en el capítulo siguiente volveré en detalle sobre esto. A modo de avance, se puede adherir a la idea de lograr una psicología científica, pero que no pretenda dar cuenta de la complejidad del comportamiento intencional sino que se ocupe de los procesos



internos que participan en las capacidades cognitivas y para ello no hace falta postular estados internos que son acerca del mundo, sino tan sólo estados que porten información interna al sistema de manera que éste pueda ejecutar sus funciones.

Lo que me interesa remarcar aquí son dos cuestiones. En primer lugar y con respecto a la tesis (9), muchos filósofos de la psicología y científicos cognitivos han argumentado que así como la física no tiene por qué incorporar o partir de la física de sentido común, la psicología tampoco debe hacerlo de la PSC dado que no hay ninguna razón de peso acerca de por qué una ciencia debería incorporar nociones de sentido común. La psicología, como toda ciencia, va desarrollando su marco conceptual en función de sus necesidades teóricas y la evidencia empírica. Si resultara que las nociones intencionales de la PSC no tuvieran ninguna correspondencia con las nociones psicológicas eso no significaría que la psicología dejaría de ser psicología (quizá, para pasar a ser neurofisiología!). Sólo quienes asimilan lenguaje psicológico con lenguaje intencional pueden sacar tan descabellada conclusión.<sup>28</sup>

Tal como hemos visto en el capítulo 1, la psicología en tanto que se mueve en el nivel de procesamiento de la información y sus explicaciones son en términos de operaciones computacionales sobre representaciones, estudia los niveles abstractos del cerebro. Este nivel intermedio de estudio, que es el ámbito de lo mental, no puede asimilarse sin más al nivel superior de sentido común o al inferior de los procesos físicos (aunque, por supuesto, sería deseable una futura integración, no asimilación, de los niveles). El nivel superior de sentido común que es donde las explicaciones intencionales podrían tener su lugar, es un nivel que requiere un abordaje interdisciplinario. Para comprender lo que hace un sistema cognitivo en su totalidad (por ejemplo, la persona) se requiere de un enfoque que tome en cuenta, entre otras cosas, las prácticas comunicativas y, que yo sepa, esto va mucho más allá de lo que está al alcance explicativo de la psicología cognitiva. Recuérdese que si Marr (1982) está en lo correcto acerca de las teorías psicológicas que se mueven en el ámbito del

---

<sup>28</sup> Este tipo de argumentos es moneda corriente en el ámbito filosófico. Se suele sostener que si no hay lenguaje intencional entonces no se está hablando acerca de la mente y por ende se estará haciendo otra cosa pero no psicología. Véase tan sólo como un ejemplo, de los múltiples disponibles, Rey (2003). En el fondo, el problema es la asimilación terminológica de distintos fenómenos. En la literatura de la filosofía de la psicología todo parece ser lo mismo: "mental", "cognitivo", "psicológico", "intencional" y "semántico".

procesamiento de información, los niveles explicativos abarcan descripciones formales y semánticas de lo que hace un sistema, descripciones de los algoritmos que ejecutan funciones y descripciones de la implementación física. En este sentido, ningún nivel apela al lenguaje intencional (volveré sobre esto en el capítulo siguiente).

En segundo lugar y con respecto a la tesis (10), hemos visto en el capítulo anterior que el PNC sostiene que para que la psicología sea científica tiene que tener leyes intencionales *ceteris paribus* que en el caso de la TDA son causales y en el caso de la teleosemántica no lo son. Para ambos casos se requiere de una teoría de cómo se determina el contenido intencional que figura en esas leyes. En lo que sigue mostraré que hay dos caminos, alternativos a los del PNC, que se pueden seguir para obtener una psicología científica sin que haya leyes intencionales. Se puede optar por la idea de que no es necesario que haya leyes psicológicas especiales para que la psicología sea considerada una ciencia (y en consecuencia no habría leyes psicológicas intencionales/causales) o se puede optar por la idea de que podría haber leyes psicológicas, pero éstas no serían causales al estilo del modelo de cobertura legal/causal que subyace a la TDA ni serían intencionales. En ambos casos, no se requiere del desarrollo de una semántica intencional porque no habría leyes intencionales en psicología.

Según Fodor, las leyes intencionales de la psicología científica son leyes *ceteris paribus* (al igual que las leyes de todas las ciencias especiales) que admiten excepciones, a diferencia de las leyes estrictas de la física que no tienen excepciones. Las generalizaciones típicamente intencionales tienen la forma: "Si ud. quiere \_\_\_\_, y cree que no puede \_\_\_\_ a menos que haga \_\_\_\_ entonces, *ceteris paribus*, ud. intentará hacer \_\_\_\_" (Fodor 1994, p. 4). Otra manera de expresar la forma de estas leyes es: 'los eventos Fs causan eventos Gs *ceteris paribus*'. Fodor sostiene que las afirmaciones singulares causales tienen que estar subsumidas por leyes causales. Así el modelo causal al que se apela es el de la subsunción bajo leyes causales. Y lo que se sostiene es que la ocurrencia de F es nomológicamente suficiente para la ocurrencia de G. Ahora bien, las leyes intencionales están implementadas en mecanismos computacionales. Típicamente, los mecanismos

que implementan las leyes de las ciencias están especificados en el vocabulario de alguna otra ciencia de nivel inferior (Cf. Fodor 1994, p. 8).

Schiffer (1991) duda que haya leyes *ceteris paribus*. Todo aquel que sostenga que puede haber este tipo de leyes tiene que hacer dos cosas. En primer lugar, tiene que explicar cuál es la naturaleza de la proposición expresada por 'Los Fs causan los Gs, a igualdad de condiciones'. Y en segundo lugar, mostrar que la verdad de las proposiciones expresadas por las oraciones *ceteris paribus* determinan leyes (porque no cualquier oración verdadera determina una ley) y que además, sean psicológicas. Con respecto a lo primero, Schiffer acepta que las oraciones *ceteris paribus* expresan proposiciones verdaderas pero, con respecto a lo segundo, éstas no constituyen leyes psicológicas. No son psicológicas porque las leyes que requieren las oraciones verdaderas *ceteris paribus* están expresadas en términos no intencionales y no son leyes porque no pueden usarse como premisas de una explicación por cobertura legal.<sup>29</sup>

Sin embargo, del hecho de que es muy dudoso que haya leyes intencionales causales *ceteris paribus*, no se sigue que no se pueda obtener una psicología científica. En consonancia con la crítica a las leyes causales intencionales, Schiffer (1991, p. 16) afirma lo siguiente:

Quando *yo* leo [textos de] biología, me cuesta mucho encontrar algo que parezca una explicación que invoque leyes, y creo que sé por qué. Supongamos que Ud. inventó una ratonera que se activa con un resorte y tuvo que explicar *cómo funcionaba*. Ud. diría que, cuando la máquina funciona, es porque un ratón mordisquea el queso puesto en un mecanismo disparador; el movimiento causado por el mordisqueo dispara una reja acoplada a un resorte estirado; etc. Pero no mencionaría ninguna ley. Quizá si continuara la explicación con una cadena causal lo suficientemente larga,

---

<sup>29</sup> Según Schiffer, el enunciado *ceteris paribus* se despliega de la siguiente manera (M está por F y B por G): "Los Ms causan Bs si para cada una de las "suficientemente muchas" realizaciones D de M hay una condición del mismo nivel C tal que D-&-C es causalmente suficiente de una manera no superflua para un evento B" (p. 7). Este bicondicional da el significado de las oraciones *ceteris paribus*. Dado que hay instancias de sustitución verdaderas del lado derecho del bicondicional, las oraciones *ceteris paribus* expresan proposiciones verdaderas. Pero las leyes aludidas por estas oraciones *ceteris paribus* dicen que una condición no-física D-&-C es nomológicamente suficiente para la ocurrencia de B. Por otro lado, ninguna instancia de sustitución del lado derecho del bicondicional es una ley de ningún tipo.

llegaría a leyes; pero serían leyes de la física, no leyes acerca de la teoría de las ratoneras. De la misma manera, la mayor parte de la biología se ocupa de explicar cómo funcionan varios mecanismos- piense en la explicación de la fotosíntesis- y parece que esas explicaciones no invocan *leyes* biológicas, estrictas o *ceteris paribus*.

La cita se refiere a la biología pero es muy probable que la mayoría de las ciencias especiales, incluida por supuesto la psicología, se ocupen de explicar cómo funciona un mecanismo y en ese sentido no requieran apelar a leyes especiales causales. El nivel donde hay que buscar las leyes causales es el nivel físico, por ende, es muy probable que las ciencias especiales no tengan leyes propias, y como consecuencia para la psicología, ésta no tendrá leyes intencionales ni causales. En todo caso las explicaciones de los mecanismos cognitivos apelarán, en última instancia, a leyes causales de la física.

Si, por otro lado, uno cree que la psicología puede tener leyes especiales pero que no serán leyes causales/intencionales, una alternativa al modelo de subsunción bajo leyes causales (Hempel y Oppenheim 1948) es la 'explicación funcional' propuesta por Cummins (1975, 1983). Cummins afirma que las explicaciones psicológicas se ocupan de explicar ciertas propiedades de un sistema y en este sentido son 'teorías de propiedades'. El objetivo de una teoría de propiedades es responder la pregunta: ¿en qué consiste para el sistema S tener la propiedad P (o más precisamente, instanciar la propiedad P)? Para ello se apela a las propiedades de los componentes de S y cómo están organizados. La idea fundamental es explicar una capacidad cognitiva compleja descomponiendo el sistema en subcapacidades más simples y menos sofisticadas y problemáticas que la capacidad que se quiere explicar tal que la manifestación de las capacidades analizadoras (especificables en un programa o un diagrama de flujo) equivale a la manifestación de la capacidad analizada.

Las capacidades de procesamiento de la información son propiedades disposicionales complejas y su análisis se denomina 'análisis funcional' (puesto que la capacidad que se analiza es una 'función' del sistema).<sup>30</sup> El primer paso del

---

<sup>30</sup> La diferencia fundamental con el enfoque teleosemántico con respecto a la noción de "función" es que para Cummins las funciones son disposiciones mientras que para Millikan, tal como hemos

análisis funcional consiste en el *análisis* de la capacidad cognitiva P y el segundo en dar cuenta de la *instanciación* en S del análisis de P. De manera que si S posee los componentes organizados de la manera especificada en el análisis, entonces S posee P. La explicación funcional concluye cuando el sistema S ejecuta el programa especificado por el análisis (S tiene que tener la estructura necesaria para cumplir las instrucciones del programa). La teoría de propiedades no requiere leyes causales como sí las requieren las 'teorías de transición' (que son teorías de subsunción bajo leyes causales) que intentan responder ¿por qué el sistema cambia del estado S<sub>1</sub> a S<sub>2</sub>? En este caso se explican los cambios como efectos de causas previas y eso requiere apelar a leyes causales. En cambio, en la teoría de propiedades se utilizan leyes (de instanciación y composición)<sup>31</sup> que no son causales, esto es, no se explica la posesión de P como efecto de alguna causa.<sup>32</sup>

Todo el peso explicativo de la explicación funcional radica, en última instancia, en la especificación de las capacidades analizadoras en el programa. Cuanto más sofisticado es el programa (esto es, cuanto más grande es la brecha entre la capacidad analizada y las capacidades analizadoras) más sentido tiene acudir a la explicación funcional. La sofisticación del programa depende de que las capacidades analizadoras sean menos sofisticadas y diferentes en tipo que la analizada. Si la brecha es pequeña, el análisis funcional pierde en capacidad explicativa. La idea es que cuanto más compleja sea la organización de los componentes de un sistema, la base física en la que se realiza deja de ser especial

---

visto, se definen en virtud de su historia evolutiva. Para una comparación entre los enfoques de ambos, véase Millikan (2002).

<sup>31</sup> Una ley de instanciación ("Cualquier cosa que tenga los componentes C<sub>1</sub>...C<sub>n</sub> organizados de manera O (teniendo un análisis [C<sub>1</sub>...C<sub>n</sub>, O] ) tiene la propiedad P", Cummins, 1983, p. 17) es una afirmación legaliforme que especifica cómo una propiedad P se instancia en un tipo especificado de sistema S. Una ley de composición ("S tiene el análisis [C<sub>1</sub>...C<sub>n</sub>,O]", Cummins, 1983, p. 17) es una afirmación legaliforme que especifica el análisis de un tipo de sistema, (por ejemplo, el modelo de la doble hélice del ADN o la composición molecular del agua). Las leyes de instanciación no son causales porque para que lo sean debe aparecer el par causa-efecto en sus instancias. Pero, requieren de explicación, esto es, se derivan de 'atribuciones nómicas' que son afirmaciones legaliformes que expresan que todo X tiene cierta propiedad P (por ejemplo, la ley de inercia y de la gravedad).

<sup>32</sup> Cuando la propiedad disposicional P se manifiesta, hay transición de estados en el sistema, por ende hay una teoría de transición que especifica las leyes de transición de la disposición P y de las disposiciones que componen la disposición P, en este sentido una teoría de transición especifica los *explananda* para una teoría de propiedades, no dice qué es para S tener la propiedad disposicional P. Cf. Cummins (1983), cap. 1.

para la capacidad analizada. Y esto es lo que permite agrupar sistemas con diferencias físicas que poseen las mismas capacidades. Es en este sentido, que el punto central de la explicación funcional es que los fenómenos psicológicos no se explican subsumiéndolos bajo leyes causales sino dando cuenta de ellos como manifestaciones de capacidades explicadas por análisis.<sup>33</sup>

Ahora bien, esta explicación no está pensada exclusivamente para dar cuenta de capacidades cognitivas. Es el tipo de explicación que se utiliza para dar cuenta de la producción de una cadena de montaje, la que aparece en los diagramas de esquemas en electrónica, en los programas de computación y los diagramas de flujo, también en las explicaciones en biológica. Sin embargo, es cierto que su uso en psicología es amplio y fructífero por lo dicho anteriormente, esto es, la posibilidad de dar cuenta de cómo un mecanismo como el cerebro, compuesto de partes organizadas de cierta manera (o, en el nivel computacional, cómo un mecanismo que almacena ceros y unos) puede dar lugar a capacidades cognitivas complejas.

No obstante, el análisis funcional, en tanto que es posible aplicarlo en diferentes disciplinas que no apelan a contenidos intencionales, no está comprometido con que las capacidades analizadoras o la capacidad analizada sean intencionales. El caso psicológico no es una excepción. Tal como hemos visto en el capítulo 2, los sistemas cognitivos instancian funciones cognitivas cuyos argumentos y valores están relacionados de manera tal que dados ciertos inputs, el output es convincente a la luz de aquellos.<sup>34</sup> Así, los sistemas cognitivos son vistos como mecanismos inferenciales, de manera que, los inputs, los estados internos y

<sup>33</sup> Dado que la explicación nomológica deductiva [N-D] no requiere leyes causales sino tan sólo leyes, se podría pensar que se le puede dar una forma N-D a la explicación funcional. Sin embargo, hay consideraciones generales y particulares para pensar que las leyes a las que apela la explicación N-D son causales. Por un lado, la explicación N-D es una expresión de la comprensión empirista de la causalidad y la explicación causal (véase Collins 1987, cap. 7). Por otro lado, según Cummins (1983, cap. 1) las teorías de transición subsumen eventos particulares bajo leyes (de ahí que son explicaciones N-D), estas leyes tienen que ser causales (subsumen pares causa-efecto) porque si no, estas teorías no serían explicativas (dado que su objeto es explicar cambios como efectos), de ahí que las leyes de las explicaciones N-D para fenómenos psicológicos sean causales. Transformar una explicación funcional en una N-D oscurecería la diferencia entre explicar cambios y explicar propiedades.

<sup>34</sup> "Convincente" o "persuasivo" serían las traducciones más aproximadas para el término inglés *cogent*. Ver nota 5 del capítulo 2.

los outputs del sistema tienen una interpretación proposicional y en ese sentido, están descriptos en términos semánticos.<sup>35</sup>

Sin embargo, esta interpretación no tiene por qué aludir a estados intencionales o a inputs y outputs descriptos en términos de las propiedades distales de los objetos del mundo externo. Gran parte del procesamiento de la información se lleva a cabo por mecanismos subpersonales en los cuales no intervienen estados de actitud proposicional y cuyos inputs y outputs son descriptos en términos que no aluden a objetos o propiedades distales en el mundo exterior al individuo. Y dado que, como se verá en detalle en el capítulo siguiente, la psicología se ocupa de estos mecanismos subpersonales, las generalizaciones que puedan establecerse acerca de estos sistemas no apelan a términos intencionales. Esto quedará mucho más claro en el capítulo 5.

Para concluir con este punto, quisiera hacer notar que la explicación funcional está en consonancia con la idea de que no hace falta leyes causales especiales para la explicación psicológica. La diferencia es que, esta última apela en última instancia al nivel físico (leyes de la física), mientras que la explicación funcional está pensada para un nivel más abstracto que es el computacional (leyes computacionales). Ambas acuerdan en lo siguiente: si de lo que se trata es de dar cuenta de por qué un mecanismo tiene cierta capacidad, entonces no hace falta apelar a leyes *causales* en el nivel computacional, sin embargo, si de lo que se trata es de explicar la transición de estados en un sistema computacional, entonces ahí sí se baja al nivel físico para dar cuenta del paso de la causa al efecto.<sup>36</sup> Cummins cree que lo que convierte en explicativa a una explicación psicológica no es esto último (una teoría de transición), sino lo primero (una teoría de propiedades). Así, tanto la propuesta de Schiffer como la de Cummins acuerdan en que la explicación

<sup>35</sup> Ver nota 6 del capítulo 2.

<sup>36</sup> Millikan (1993b) estaría de acuerdo con esta descripción de la explicación psicológica. Adhiere a la explicación funcional de Cummins (con la salvedad mencionada en la nota 30) y en este sentido a la idea de que para encontrar leyes causales la explicación llegaría hasta la estructura física del sistema y las condiciones físicas para sus operaciones. Los pasos de la explicación funcional permitirían conectar la explicación Normal teleológica con la explicación causal que apela a leyes físicas. El punto de mi disputa con la explicación psicológica de la teleosemántica (y el otro aspecto por el cual, a mi entender, se diferencia de la explicación funcional) es que la explicación funcional

psicológica no requiere de leyes causales intencionales especiales.<sup>37</sup> De manera que, no haría falta una teoría del contenido intencional para la psicología cognitiva.

Las tesis (9) y (10) quedarían reformuladas así:

7- *Psicología científica*: se puede lograr una psicología científica.

En relación con la tesis (3) del naturalismo metafilosófico, tal como se mencionó en el capítulo anterior, si bien las tesis (1) del naturalismo y (2) del naturalismo científicista la suponen, sostener (3) no implica la adhesión a las tesis (1) y (2). No veo ninguna objeción de principio a mantener que la filosofía puede hacer uso de los datos de las disciplinas científicas, siempre que se considere a la ciencia en el sentido amplio de abarcar disciplinas como las humanas, o que algunas de sus tesis puedan sean empíricas (es esta una de las razones por las cuales la primera estrategia *a priori* de ataque al PNC, que hemos visto en el §1.1, no parecía convincente). Se puede tener esta actitud naturalista sin comprometerse con la afirmación reduccionista involucrada en las tesis (1) y (2). De manera que la tesis (3) quedaría reformulada así:

2- *Naturalismo metafilosófico*: la filosofía de la mente se desarrolla en conformidad con, o hace uso de, la ciencia corriente para contribuir a la explicación de los diferentes aspectos del ámbito humano.

Por otro lado, tal como (3) está formulada, el PNC parece asumir que la mente humana no es algo natural *per se*. De aquí la idea de explicar las propiedades mentales, en particular la intencionalidad, en términos de propiedades físicas. Esto es porque, por un lado, se asimila naturalismo con científicismo, y por el otro, se le exige a las ciencias naturales más de lo que pueden dar. Esto lleva a una especie de tensión entre un dualismo metodológico y una incorporación a las ciencias naturales de elementos extraños a las mismas. Por un lado, se trata a la mente como algo distinto

---

no se compromete con contenidos intencionales, y en este sentido, las leyes computacionales no tienen por qué implementar leyes intencionales (aunque no-causales).

<sup>37</sup> Este punto le debe en mucho a una discusión con Diana Pérez.



de un objeto natural y, por el otro, se intenta que este polo del dualismo se inserte en la ciencia y que ésta explique, en su totalidad, fenómenos como la intencionalidad.

A diferencia del PNC, el PNA trata a su objeto de estudio, la mente, como parte del mundo natural. Así, la tarea de la filosofía de la psicología no consiste en elaborar propuestas para conectar lo mental con lo físico sino en dar cuenta (junto con el resto de las disciplinas humanas) de cómo los humanos nos arreglamos para interactuar con el entorno y otros humanos. El mundo no es ahora el mundo estrecho de la física ni tampoco el mundo de las ciencias naturales. Tenemos que entender el “mundo” de una manera diferente de cómo lo entienden las teorías científicas del contenido mental. El mundo contiene tanto los objetos físicos como nuestras prácticas comunicativas que están ancladas en nuestras “formas de vida” que, a su vez, están ancladas en nuestros sistemas senso-perceptivos y en ciertas regularidades.

Dado que la intencionalidad es una cuestión de interpretación de los estados internos de una persona en función de las prácticas comunicativas y éstas se basan en capacidades y acciones de las personas que están ancladas en el mundo natural, no hay necesidad de un naturalismo científico. Dicho de otra manera, si se abandona la idea de que el naturalismo es igual a científico y que hay un contenido a naturalizar, entonces se puede sostener que lo que pensamos, creemos, y demás, forma parte del mundo natural, y podemos dejar a la psicología que se ocupe de los procesos cognitivos internos y apelar al resto de las disciplinas para dar cuenta de cuestiones como la intencionalidad y, en general, del comportamiento humano. Es en este sentido que las tesis básicas del PNA establecen el marco adecuado para el desarrollo de una teoría de la atribución intencional.

A continuación aparecen en conjunto las tesis del PNA (los números originales de las tesis del PNC aparecen entre paréntesis para que se puedan apreciar los cambios):

1-(1/2) *Naturalismo*: los pensamientos, las creencias y similares de las personas forman parte del mundo natural porque están anclados en prácticas comunicativas que a su vez, se anclan en “formas de vida”, las que a su vez están ancladas en sistemas senso-perceptivos y ciertas regularidades físicas y comportamentales.

2-(3) *Naturalismo metafilosófico*: la filosofía de la mente se desarrolla en conformidad con, o hace uso de, la ciencia corriente para contribuir a la explicación de los diferentes aspectos del ámbito humano.

3-(4/5) *Realismo intencional*: existen estados internos en las personas que son aquellos a los que se alude en la atribución intencional por medio de oraciones de AP. La naturaleza de estos estados refleja la naturaleza lingüística de las oraciones de AP.

4-(6) *Externalismo*: el contenido de los estados intencionales de una persona está determinado por sus relaciones con el mundo.

5-(7) *Indeterminación del contenido*: no se requieren hechos decisivos con respecto a cuál es el contenido de un cierto estado mental de una persona.

6-(8) *Molecularismo del contenido*: la identidad del contenido de un estado intencional depende de los estados intencionales asociados en contextos particulares.

7-(9/10) *Psicología científica*: se puede lograr una psicología científica.<sup>38</sup>

---

<sup>38</sup> Está claro que las tesis del PNA requerirían de una mayor elaboración, pero llenar los detalles y profundizar en ciertos aspectos es justamente lo que se necesita para una teoría sustancial de la atribución intencional y, tal como he mencionado en reiteradas oportunidades, mi objetivo aquí no es ese sino tan sólo esbozar el marco que permita una alternativa al PNC.

## 5

## Representaciones sin semántica intencional

En el capítulo anterior dije que la concepción del PNC es infructuosa porque al asimilar el fenómeno de la intencionalidad al de representación dejaba sin explicar cuestiones fundamentales y propias de cada fenómeno. Por el lado de la intencionalidad, hemos visto cómo el PNC simplifica este fenómeno a una relación entre propiedades dejando de lado elementos que tendrían que participar en dicha relación dado que son indispensables para comprender cómo las personas se las arreglan para pensar acerca del mundo. Por el lado de las representaciones, dije que considerarlas como portadoras de propiedades intencionales oscurecía su rol explicativo en las teorías cognitivas.

En el capítulo 2 (§1.2) vimos que un abordaje del problema de las representaciones debía cumplir con los requisitos: (a) explicativo, (b) implementacional y (f) integración cognitiva. El requisito explicativo establece que una teoría acerca de las representaciones debe sustentar, o al menos, ser compatible con el rol explicativo que las ciencias cognitivas le otorgan a las representaciones. A lo largo de este capítulo veremos que las teorías psicológicas no le adscriben a las representaciones propiedades intencionales, de manera que el rol explicativo de las mismas no es el de dar cuenta del fenómeno de la intencionalidad, como cree el PNC, sino el de dar cuenta del funcionamiento de los mecanismos cognitivos internos subpersonales. En este sentido, el PNC, no cumpliría con, al menos, uno de los tres requisitos fundamentales de una teoría de la representación.

Para mostrar esto me centraré en las mismas teorías cognitivas que los filósofos de la psicología suelen analizar en la literatura y que he presentado en el capítulo 1 (§1.5.1). Estas teorías son acerca de la capacidad lingüística y visual, y sus precursores fueron Chomsky y Marr, respectivamente. Si bien, a lo largo del

capítulo diré varias cosas con respecto a la teoría marriana, me concentraré en la teoría chosmkiana. Para aquellos que tengan ciertos escrúpulos en considerar a ambos desarrollos como teorías paradigmáticas de la psicología cognitiva, es importante remarcar, tal como se dijo en el capítulo 1 (§1.4 y §1.5), que ambas responden a las tesis básicas de la psicología cognitiva enunciadas en ese mismo capítulo y en ese sentido, integran la psicología cognitiva de manera legítima. Por otro lado, si aún así los escrúpulos no se diluyen, quizá por prejuicios acerca de lo que los científicos cognitivos de diferentes corrientes consideran que son las teorías paradigmáticas, también es importante remarcar que son las teorías de las cuales los filósofos de la psicología se han ocupado con mayor detalle y profundización. De manera que si logro mostrar que las representaciones postuladas en esas teorías no requieren de propiedades intencionales, el PNC será refutado en su propio terreno.

Como el objetivo es mostrar que la intencionalidad no tiene lugar en las teorías psicológicas porque éstas se ocupan de mecanismos y estados cognitivos no-intencionales, se requiere mostrar, en primer lugar, que hay alguna distinción entre los estados intencionales de AP y los estados de los mecanismos cognitivos que procesan información y, en segundo lugar, que esos estados que procesan información no son intencionales. Con respecto a lo primero, en la literatura filosófica se han propuesto, al menos, dos distinciones. Una es la distinción entre estados (y mecanismos) doxásticos y subdoxásticos, y la otra, la distinción entre estados (y mecanismos) personales y subpersonales (hemos visto algo sobre esta última, en particular, en el capítulo 2 (§1.1)). En el §1, analizaré críticamente la primera con el objetivo de mostrar que no logra establecer la distinción en cuestión y que, en última instancia requiere de la distinción personal-subpersonal (si es que no colapsa con ella). No obstante, en el transcurso de esta evaluación veremos que, a pesar de no lograr establecer una diferencia de principio entre los estados intencionales y los estados de procesamiento de la información, esclarece algunas de las características de ambos.

La propuesta para la comprensión del rol de la noción de representación en las teorías acerca de las capacidades lingüística y visual se basa entonces en la distinción entre el nivel personal y subpersonal. En el §2 expondré con mayor

detalle el funcionamiento del mecanismo lingüístico e, introduciré la noción de 'contenido interno'. Mostraré que el rol explicativo de las representaciones mentales, que intervienen en los mecanismos psicológicos subpersonales, es comprensible a la luz de la idea de que los estados subpersonales poseen un contenido que no es intencional sino que es interno a los subsistemas de procesamiento de la información.

Dado que la noción de contenido interno puede dar lugar a confusiones en relación con otra noción de contenido que se ha propuesto en la literatura filosófica para dar cuenta del tipo de contenido que interviene en la explicación psicológica, en el §3 intentaré distinguir la noción de contenido interno de la de 'contenido estrecho'. En el §4, ampliaré la idea de que son los estados subdoxásticos subpersonales, de la manera en que resultarán especificados a partir de la noción de contenido interno, aquellos de los cuales la psicología cognitiva está en condiciones de dar cuenta por sí misma, puesto que son estos estados los que intervienen en las explicaciones psicológicas cuyo objetivo es dar cuenta de las capacidades cognitivas.

## 1. Estados subdoxásticos y estados doxásticos

Fue Stich quien distinguió entre estados intencionales como los de creencia y estados que "juegan un rol en la historia causal próxima de las creencias, aunque no son creencias en sí mismos" (1978, p. 499), y denominó a estos últimos Estados Subdoxásticos (ES). Los ESs son estados psicológicos que almacenan información acerca de, por ejemplo, las reglas de la gramática<sup>1</sup> que llevan a juicios de gramaticalidad o, los rasgos de la imagen retinal que dan lugar a juicios perceptuales. En ambos casos hay mecanismos que median entre el estímulo

---

<sup>1</sup> Chomsky (1995a, p. 17) afirma que en el programa minimalista no hay reglas para formar construcciones gramaticales (como en la gramática generativa de etapas anteriores), sino, tal como veremos en §2, un conjunto de principios y parámetros. De todas formas, el ejemplo de Stich se aplica igual.

(auditivo o retinal) y la creencia, cuya tarea consiste en utilizar la información almacenada en sus estados para dar lugar (quizá junto con otros mecanismos cognitivos) a la creencia correspondiente. Estos estados forman parte del proceso de formación de creencias sin ser ellos mismos creencias. Evans (1982) también distingue entre estados informacionales y creencias, siendo los primeros estados de sistemas que procesan y almacenan información, cuya caracterización es independiente de la noción de creencia.<sup>2</sup>

Se han ofrecido tres criterios para distinguir los ESs de los estados intencionales: (a) la accesibilidad a la conciencia (Stich 1978), (b) la integración inferencial (Stich 1978), y (c) la conceptualización (Davies 1986, 1989). En lo que sigue analizaré estos criterios, a través de la discusión de los argumentos de Davies (1989) en contra de (a) y (b) y a favor de (c), con la intención de mostrar las ventajas y desventajas que poseen para una distinción de principio. Cabe aclarar que la perspectiva general desde la cual se ofrecen los criterios es a partir del análisis del concepto ordinario preteórico de creencia.<sup>3</sup> Es este análisis *a priori* el que pretende justificar la 'distinción intuitiva' entre los estados intencionales (como los de creencia) y los ESs con el objetivo de mostrar que la distinción tiene interés psicológico.<sup>4</sup> El análisis del concepto ordinario de creencia muestra que el individuo que las posee tiene conciencia del contenido de sus creencias, que éstas están inferencialmente integradas, y que el contenido es conceptual. Mientras que con respecto a los ESs, el individuo no tiene acceso conciente a sus contenidos, están aislados inferencialmente y su contenido es no conceptual. Veamos en detalle estos rasgos.

---

<sup>2</sup> Así, "que el sujeto *esté* en un estado informacional es independiente de si cree o no que el estado es verídico" (Evans 1982, p. 123).

<sup>3</sup> El rasgo (a) es producto de motivaciones *a priori* y empíricas. Propuesto inicialmente por Chomsky, (a) está motivado por consideraciones empíricas, en particular y tal como veremos más adelante, por argumentos a la mejor explicación psicológica. En las versiones de Stich (1978) y Davies (1989), la motivación es *a priori*.

<sup>4</sup> Ni en Stich (1978) ni en Davies (1989) se especifica cuál sería el interés para la psicología (más allá de que son estados que jugarían un rol causal en la producción de las creencias). En el §4 intento mostrar que no sólo la distinción interesa a la psicología sino que son los ESs aquellos de los cuales la psicología está en condiciones de dar cuenta.

### 1.1. (a) *Accesibilidad a la conciencia*

A partir de los primeros desarrollos de la teoría chomskiana, el rasgo de inaccesibilidad a la conciencia está ligado a los mecanismos y estados cognitivos del lenguaje-I y desde entonces Chomsky se ha encargado de defender en (casi) todos sus escritos la idea de que éstos no sólo no son *efectivamente* accesibles a la conciencia sino que ni siquiera son *potencialmente* accesibles a la conciencia (Chomsky 1965, p. 8; 1980a, pp. 137-42; 1986, p. 260, pp. 269-71; 1994a, p. 93 y ss.). El sentido de “conciencia” al que se alude es en términos de poder reportar, asentir a, o reflexionar sobre algo. Los hablantes no pueden en general reportar sus estados lingüísticos y cuando lo intentan, la mayoría de las veces suelen estar equivocados, de manera que los contenidos no están abiertos a la reflexión. Además, los hablantes no reconocen ni asienten cuando son informados acerca de las representaciones mentales involucradas en el uso del lenguaje. Esto muestra que no hay acceso privilegiado por introspección a los mismos.

Todo esto es evidencia a favor de que los estados lingüísticos son inconcientes, pero en realidad el argumento fundamental en defensa de que hay estados y mecanismos inconcientes tiene la forma de los argumentos a la mejor explicación. Dado cierto fenómeno a explicar y dado que de todas las explicaciones en competencia, la explicación que postula estados y procesos inconcientes es la mejor, según ciertos criterios de elección entre explicaciones, es probable que esta explicación sea la adecuada. El criterio de elección al que se suele aludir es el poder explicativo de la teoría y no parece haber ninguna explicación que no postule un mecanismo inconciente, que pueda dar cuenta de los numerosos hechos que explica una teoría que los postula (Chomsky 1980a, pp.139-40).<sup>5</sup> Otro criterio al que se suele apelar es el de la familiaridad o concordancia con tipos de explicaciones aceptadas, así la explicación que postula estados y procesos inconcientes concuerda con el tipo de explicación habitual de las ciencias cognitivas que postula que gran parte del procesamiento de la información ocurre a nivel inconciente.

---

<sup>5</sup> Tales como la dependencia referencial, ver nota 30.

Desde el punto de vista del análisis del concepto ordinario de creencia la conclusión es la misma: los ESs son inconcientes. Stich (1978), al igual que Chomsky, analiza la accesibilidad a la conciencia de los contenidos de las creencias en términos de la posibilidad de reportarlos. Un individuo puede tanto decir si cree que  $p$  como, si cree que  $p$ , reportar que  $p$  es el caso. Y esto es así porque en los casos típicos de creencia, el sujeto es conciente de los contenidos en el sentido de tener la experiencia conciente de “tener la creencia ocurrente de que  $p$ ” (p. 504). En cambio, los contenidos de los ESs no son accesibles a la conciencia en el sentido de que el individuo no puede reportarlos ni puede asentir a los mismos.

Como se puede notar tanto Chomsky como Stich consideran el fenómeno de la conciencia en términos de poder reportar y asentir.<sup>6</sup> Según Davies (1989), esta noción preteórica de conciencia no es adecuada como criterio de principio para la delimitación entre los ESs y los estados intencionales por tres razones:

- 1- Hay creencias inconcientes, de manera que ser inconciente no es condición suficiente para ser un ES.
- 2- Hay aspectos de la red inferencial de creencias que son inconcientes.
- 3- El ámbito de lo conciente es más amplio que el de los estados intencionales.

De las tres razones, parece que 1 es la de más peso. El argumento que subyace a 1 es que tenemos una noción de creencia inconciente, si no se quiere que colapse con la noción de ES habría que poner restricciones. Se podría restringir la noción de creencia inconciente utilizando una variante del principio searliano de que todo estado intencional es inconciente sólo en el sentido de que es ‘en principio accesible a la conciencia’ salvo un mecanismo de bloqueo (Searle 1990). Pero esto no sería suficiente porque se corre el riesgo de clasificar casos paradigmáticos de

---

<sup>6</sup> Aunque como se puede advertir en el párrafo anterior, Stich también habla de ‘experiencia conciente’ atribuyéndole así a los estados intencionales además de propiedades de la conciencia de acceso, propiedades de la conciencia fenoménica (Véase la distinción en Block 1995a). Dado que el tipo de conciencia que está en juego en los estados cognitivos es la de acceso y dado que es discutible que los estados como los de creencia tengan propiedades fenoménicas (Cf. Davies 1995 para ambos puntos), es conveniente tomar la noción de conciencia en el sentido de acceso.



ES como estados no accesibles a la conciencia por un mecanismo de bloqueo, con lo cual, la noción de ES colapsa otra vez con la de creencia inconciente. Por ende, habría que restringir aún más este principio, pero parece que cualquier restricción que se piense, probablemente constituiría un criterio independiente para fundar la distinción.

Según mi opinión, el argumento de Davies descansa en la aceptación de que puede haber algo así como creencias inconcientes. Creo que si se niega esta posibilidad, el criterio de accesibilidad a la conciencia sería un criterio delimitador de principio. No obstante, esta estrategia no funciona porque parece difícil negar que tenemos una noción de *sentido común* de creencia inconciente. Sin embargo, se puede responder a esto que lo que está en juego aquí no son nociones de sentido común (difícilmente se tenga una noción de sentido común de ES) sino nociones *teóricas*, esto es, que participan en teorías, en particular, las psicológicas. Si esto no fuera así, no se ve por qué la distinción doxástico-subdoxástico tendría que tener importancia para las *teorías* psicológicas. La noción de mecanismos y estados de procesamiento de información inconcientes participa, al menos, de teorías bien establecidas como las del lenguaje (Chomsky) y la visión (Marr). En cambio, a mi entender, es dudoso que la noción de creencia inconciente participe de alguna teoría. Stich (1978), Searle (1990) y Davies (1989) consideran que esta noción es clave en la teoría psicoanalítica (Freud). Sin embargo, no está claro que Freud hable de creencias inconcientes. La noción controvertida no es el adjetivo 'inconciente' sino la adscripción del mismo a las creencias. No está claro que el psicoanálisis maneje la misma noción de creencia que la elaborada por la filosofía de la mente (aunque pudiera aparecer este término en la teoría psicoanalítica). No es este el lugar para un análisis de la teoría freudiana, mi intención es sólo señalar que la aceptación de la noción teórica (no la de sentido común) de creencia inconciente requeriría de una elaboración mucho más profunda que la que le dedica Davies en su crítica al criterio de accesibilidad a la conciencia.

Ahora bien, si aún a pesar de todo esto, resultara que hay una noción de creencia inconciente que juega un rol en las teorías cognitivas y si lo único que distingue las creencias de los ESs es la accesibilidad a la conciencia, es porque se

considera que su naturaleza es la misma. Si tener una creencia consiste en estar en una relación computacional con representaciones almacenadas, tener un ES consiste en tener algún tipo de relación computacional con representaciones almacenadas. Podría ser el caso que las creencias inconcientes se distinguieran de las concientes y de los ESs inconcientes por el tipo de actitud que se mantiene con el contenido. Quizá una actitud distinta de las hasta ahora consideradas. Pero, está claro que esta restricción ya no estaría anclada en la accesibilidad a la conciencia sino que se fundamentaría en un criterio independiente acerca de la naturaleza de las creencias inconcientes, y en ese sentido Davies tiene razón en que parece que cualquier intento de restringir la noción de creencia inconciente nos lleva a otro criterio delimitador.<sup>7</sup>

La razón 2 descansa en la observación de Fodor (1983, p. 85) de que un individuo puede realizar inferencias que son instancias del *modus ponens*, puede ofrecer las premisas como razones para la conclusión y sin embargo, no puede formular la regla en su generalidad. "Nuestro compromiso" con la regla de inferencia no parece un caso *prima facie* del dominio subdoxástico aunque ésta sea inconciente. Davies ensaya dos respuestas imaginarias de los defensores del criterio de accesibilidad-inaccesibilidad a la conciencia, basadas en la distinción entre estados representacionales y procesos computacionales. La primera sostiene que la accesibilidad a la conciencia es un rasgo de los estados y no de los procesos, de manera que el hecho de que haya reglas inconcientes en la red inferencial de creencias no amenaza el criterio de accesibilidad a la conciencia del contenido de las creencias.

Davies le objeta a esta defensa que de esta manera el rasgo de inaccesibilidad a la conciencia de las reglas lingüísticas (que, según Davies, muy probablemente se instancian como procesos computacionales más que como estados representacionales) no sería de peso porque habría reglas que intervienen en el ámbito intencional que también serían inconcientes y de este modo se borraría el

---

<sup>7</sup> De hecho es lo que hace Davies (1995) al refinar la noción de conciencia al punto de hacerla depender del criterio de conceptualización, para luego tomar a este último como el criterio sobre el cual descansa la distinción entre ESs y estados intencionales.

criterio de accesibilidad/inaccesibilidad a la conciencia y con ello la distinción entre el ámbito intencional y el subdoxástico. Sin embargo, esto no es así. Si uno creyera, por ejemplo, que todo lo que hay para lo subdoxástico en la teoría chomskiana son sólo reglas, entonces sería cierto que el criterio de accesibilidad no distinguiría el ámbito subdoxástico del intencional. Pero no puede ser así porque, aunque las reglas estén instanciadas en los mecanismos de procesamiento computacional, éstas operan sobre información, por ejemplo, entendida como instancias de representaciones. Estas instancias son las que figuran en los estados cognitivos subdoxásticos, de manera que la información que extraen los algoritmos computacionales es la información que aparece en los estados computacionales subdoxásticos. La distinción subdoxástico-doxástico para los estados se sigue manteniendo más allá de que las reglas que operen sobre los ESs o los estados intencionales no sean accesibles a la conciencia.

Por otro lado, la segunda respuesta a la objeción de Davies sostiene que si el rasgo en cuestión provee un criterio claro de distinción entre estados intencionales y ESs entonces también abona una distinción entre reglas mentales en el dominio intencional y procesos computacionales en el dominio subdoxástico. Tal como se desprende de lo dicho con respecto a la primera respuesta no creo que haga falta esta distinción, pero Davies contesta acertadamente que esto depende de si el criterio provee una distinción entre estados y para mostrar que esto no es así, la razón 3 apunta directamente a los estados. Dado que como se verá, el argumento de Davies que avala 3 no parece adecuado, la razón 2 se queda sin el apoyo necesario para objetar el criterio de accesibilidad a la conciencia.

La razón 3 se basa en un caso imaginado: supongamos que cierto aspecto del carácter fenoménico de una experiencia depende, de una manera sistemática, de cierto ES, de modo que ese estado sería, en cierto sentido, accesible a la conciencia. Sin embargo, esto no convertiría a ese aspecto de la experiencia en candidato a creencia. La idea es que si lo único que distingue a los ESs de las creencias es que son inconcientes, si los ESs fueran concientes entonces deberían convertirse en creencias. Me parece que el caso imaginado no es el adecuado porque lo que pone en juego no es la conexión entre ES y las creencias sino entre las experiencias y las

creencias perceptuales. Es sabido que las experiencias perceptuales pueden no llevar a ninguna creencia o llevar a incredulidad (no dar crédito a lo que se percibe, por ejemplo, en el caso de las ilusiones ópticas, McDowell 1994a, cap. 3; Martin 1994). Si imaginamos, tal como pide Davies, que un aspecto de esa experiencia perceptual esté asociado a un ES, ese aspecto de la experiencia no sería candidato a creencia simplemente porque las experiencias son independientes de las creencias.<sup>8</sup>

### *1.2. (b) Integración inferencial*

Las creencias están relacionadas entre sí de manera que conforman una red de inferencias deductivas e inductivas potenciales, de modo que los caminos por los cuales una creencia puede dar lugar a otra son innumerables al punto de que casi cualquier creencia puede llevar a cualquier otra (Stich 1978 denomina a esto 'promiscuidad inferencial'), están al servicio de distintos propósitos (Evans 1981) y pueden ser usadas (combinadas con deseos) como razones para la acción (Evans 1982, cap. 5; Davies 1989). En cambio, los ESs no pueden combinarse ellos mismos con deseos para usarse como razones para la acción. Los ESs están aislados inferencialmente del sistema de creencias de un individuo, en el sentido de que tienen un rango limitado y empobrecido de patrones inferenciales posibles para dar lugar a creencias o ser producto de ellas. No es que no haya relaciones inferenciales entre ellos sino que las relaciones están empobrecidas, los ESs están parcialmente aislados de la red inferencial de creencias.

Por un lado, los ESs no se combinan con creencias para dar lugar a otras creencias. Un lingüista que tiene una creencia explícita acerca de una regla (o principio) puede realizar inferencias en conjunción con otras creencias, pero un hablante que sólo tiene información gramatical almacenada no puede utilizar esta

---

<sup>8</sup> La idea es que un ES de procesamiento de información puede dar lugar a una experiencia perceptual que a su vez puede dar lugar a una creencia perceptual. Este caso imaginado es acerca de experiencias perceptuales, que son siempre concientes. Si un ES es parte de una experiencia

información junto con otras creencias para dar lugar a otra creencia. El ejemplo de Stich (1978) consiste en que la creencia que si la regla  $r$  entonces Chomsky está seriamente equivocado junto con el ES que almacena  $r$ , no lleva a la creencia de que Chomsky está seriamente equivocado (la creencia de que si  $P$  entonces  $Q$  combinada con el ES que  $P$ , no da lugar a la creencia que  $Q$ , como sí lo haría si las premisas fueran ambas creencias). Por otro lado, los ESs no se combinan con creencias para dar lugar a otros ESs porque los sistemas que operan con ESs no son cognitivamente penetrables por las creencias, los deseos, y el conocimiento en general del individuo. Las relaciones inferenciales entre los ESs y las creencias son especializadas y restringidas porque los ESs son parte de subsistemas cognitivos de propósito específico y en este sentido, las relaciones inferenciales entre los ESs de distintos subsistemas también son limitadas (y quizás inexistentes, en algunos casos).

La estrategia de Davies no es argumentar directamente en contra del criterio de integración inferencial. Hacer esto último podría haber consistido en argumentar en contra de que los ESs no estén integrados inferencialmente con la red de creencias o bien mostrar que la red de creencias no está integrada inferencialmente de la manera optimista en la que se cree que está integrada (promiscuidad inferencial), o hacer ambas cosas. Davies no está interesado en atacar (b) sino que su estrategia consiste más bien en aceptarlo y sugerir que el criterio de conceptualización puede fundamentarlo y así proporcionar un criterio fuerte para la distinción entre los ESs y los intencionales. Dicho de otro modo, la razón por la que no discute (b) es que considera que (c) lo fundamenta, y en este sentido, el criterio (b) constituye una distinción pero no de principio porque depende de (c).

Tal como se mencionó antes, una estrategia para discutir (b) sería argumentar que la red de creencias no está integrada de la manera en que Stich y otros creen. Stich (1978) parece decir que las creencias están sometidas a la clausura bajo consecuencia lógica: si uno cree que  $p$  entonces  $q$ , y si cree que  $p$

---

conciente, no por ello se transforma en creencia porque las experiencias no siempre llevan a creencias. Ver §1.3 y nota 20.

entonces, llega a creer que  $q$ . La clausura no se sostiene si interviene un ES porque si uno tiene una creencia condicional  $C$  y un ES  $S$  (que es el antecedente de  $C$ ), uno no llega a creer el consecuente del condicional. Sin embargo, hay ejemplos en los que las creencias tampoco se combinan de manera de dar lugar a una nueva. Richard (1994) propone dos nociones de promiscuidad inferencial (que podrían extraerse del artículo de Stich) y muestra que la razón por la cual no habría integración inferencial entre ESs y las creencias no se debe al carácter subdoxástico de alguna de las premisas.

La primera propuesta es entender la promiscuidad inferencial en términos de que los contenidos determinan condiciones de verdad y relaciones de implicación, y si las condiciones son las adecuadas, están bajo clausura de implicación lógica. Bajo esta definición, la combinación de la creencia de que si Twain duerme entonces hay problemas junto con la creencia de que Clemens duerme no lleva a la creencia de que hay problemas, porque aunque 'Twain duerme' y 'Clemens duerme' tienen las mismas condiciones de verdad, el contenido es distinto porque son las propiedades lógicas (sintácticas) las que determinan la identidad de contenido.<sup>9</sup> En general, ocurre que los estados subdoxásticos no comparten contenidos con las creencias porque probablemente difieren en estilo de representación y dado que la identidad de contenido depende (en parte) de relaciones lógicas y éstas de la sintaxis, las interacciones entre ambos tipos de estados estarían bloqueadas.<sup>10</sup> Así:

Para que dos estados tengan un tercer estado como consecuencia lógica, tienen que tener contenidos relacionados adecuadamente. (Compare: para que dos oraciones tengan una tercera oración como consecuencia lógica, tienen que tener una sintaxis relacionada adecuadamente). Dado que las creencias y los estados subdoxásticos no comparten contenidos, no tienen consecuencias lógicas (significativas). (Compare: dado que 'Twain duerme' y 'Si Clemens duerme, habrá problemas' no están

<sup>9</sup> La opción de que el contenido sea sólo condiciones de verdad no parece ser una salida, véase Richard (1994), pp. 305-6.

<sup>10</sup> Uno puede pensar (aunque Richard no lo dice con estas palabras) que la identidad del contenido depende, en parte, de los modos de presentación entendidos como vehículos sintácticos del contenido.

adecuadamente relacionadas en contenido, no tienen consecuencias lógicas interesantes). Así, que haya relaciones inferenciales relativamente empobrecidas entre creencias y estados subdoxásticos no muestra *bajo la caracterización de promiscuidad en cuestión* que las creencias y los estados subdoxásticos no son conjuntamente promiscuos. Dado que no comparten contenidos no *necesitan* interactuar inferencialmente... para ser conjuntamente promiscuos, puesto que un conjunto de estados es promiscuo cuando sus miembros *que tienen contenidos inferencialmente relacionados* se combinan para producir nuevos estados. (Richard 1994, p. 306)

Si no se acepta esta definición de promiscuidad inferencial, la segunda propuesta es entender la promiscuidad inferencial en términos de patrones de inferencias entendidos como mecanismos causales, así “[una creencia] *b* es promiscua con la colección de estados-tipo *S* siempre que haya, para muchos de los miembros *b'* de *S*, algún subconjunto “adecuado” *X* de *S* y un patrón de inferencia *I* tal que  $X, b \vdash b'$ ” (p. 307). El problema es que ninguna noción plausible de patrón inferencial impediría que los ESs sean promiscuos con las creencias. Si se entiende la noción en términos de clausura bajo producto, se podría pensar que cierto patrón de inferencia *I* proyecta un ES *S* en una creencia *C*, que a su vez es proyectada por otro patrón inferencial *I'* a otra creencia *D'*. Podría ser posible agrupar los mecanismos causales de manera de dar lugar a (o inferir) un patrón (o mecanismo) *I''* que lleve de *S* a *D'*.

Por otro lado, si la noción de patrón inferencial se entiende en términos de la distinción entre mecanismos de propósito específico y de propósito general (siguiendo la idea de Stich de que las relaciones inferenciales son especializadas puesto que los mecanismos en los que intervienen los ESs son de propósito específico),<sup>11</sup> no parece descabellado suponer que una regla general como el *modus ponens*, entendida en términos de un mecanismo causal que se aplica a las

---

<sup>11</sup> Richard observa que es difícil individuar patrones de inferencias generales o específicos porque se corre el riesgo de que no sean tan generales como para cubrir la mayoría de las creencias o que sean tan generales que incluyan ESs.

creencias, no se aplique también a los ESs.<sup>12</sup> Si contraejemplos como los de Richard resultan válidos, esto mostraría que la distinción en cuestión no puede basarse en el criterio (b).

### 1.3. (c) *Conceptualización*

La idea de que tener una creencia requiere posesión de conceptos es una de las pocas cosas en las que acuerdan los filósofos de la mente. Un individuo no puede creer que  $p$  (o tener el pensamiento que  $p$ , donde  $p$  es, por ejemplo, 'el agua es inodora') sin tener los conceptos involucrados en la proposición (AGUA e INODORA). Según Evans (1981), el que una creencia esté al servicio de distintos proyectos supone que el individuo es sensible al lugar que ocupa esa creencia en su red de creencias, esto es, es sensible a su potencial inferencial: las creencias de las cuales se infiere esa creencia y las que se pueden inferir de ella. Esto, a su vez, supone que las creencias (pensamientos, en general) son estados estructurados en el sentido de ser un complejo de habilidades, por ejemplo, creer que  $Fa$  consiste en el ejercicio de dos habilidades, la correspondiente a  $a$  y a  $F$ .

Así, ser sensible al potencial inferencial de la creencia de que  $a$  es  $F$  depende de la sensibilidad al potencial inferencial de otras creencias en las que están involucrados  $a$  y  $F$  (el pensamiento  $Fa$  sería la intersección de dos series, la de  $Fa, Fb, \dots$ , y la de  $Fa, Ga, \dots$ ). Si un sujeto tiene los pensamientos estructurados de que  $a$  es  $F$  y  $b$  es  $G$  no habría en principio ninguna barrera conceptual para tener el pensamiento de que  $a$  es  $G$  o que  $b$  es  $F$ . Es en esta capacidad de ejercitar habilidades que consiste la posesión de conceptos, de manera que tener una creencia depende de habilidades conceptuales estructuradas.

Así, si a un sujeto se le atribuye el pensamiento que  $a$  es  $F$ , entonces debe tener los recursos conceptuales para tener el pensamiento que  $a$  es  $G$ , para cualquier

---

<sup>12</sup> La idea sería que hay un mecanismo causal que cada vez que se tiene un ES  $S$  causa un estado de creencia  $C$ , que sería plausible suponer que hay una instancia de un estado representacional  $S/C$



propiedad de ser *G* de la cual tiene una concepción. Esta es la condición que llamo 'La restricción de generalidad'. (Evans 1982, p. 104)

Siguiendo la propuesta de Evans (1982, p. 104, n. 22) de que la restricción de generalidad se aplica al ámbito intencional pero no a los estados informacionales, Davies (1989) considera que esta restricción provee el criterio buscado para una distinción fundada, en el sentido de que explicita parte de lo que está involucrado en la posesión de conceptos, siendo la distinción entre dos tipos de contenido la que carga todo el peso. Los estados intencionales tienen contenido conceptual a diferencia de los ESs cuyo contenido es no-conceptual. Dado que poseer un contenido intencional (de que *a* es *F*) es *ipso facto* cumplir con la restricción de generalidad (ejercitar el dominio del concepto de *ser F*), la restricción de generalidad es, parcialmente, constitutiva de la noción de contenido conceptual.<sup>13</sup> Los conceptos son los constituyentes de los pensamientos, de manera que tener un pensamiento es *ipso facto* tener los conceptos que lo constituyen. En cambio, no se requiere que los conceptos utilizados para la atribución de contenido informacional sean poseídos por el individuo que está en ese estado. Desde esta perspectiva, ambos tipos de contenido son representacionales, en el sentido en que el mundo está representado de cierta manera, pero en el caso del no-conceptual no se ejercitan habilidades conceptuales.

Una estrategia para discutir el criterio de conceptualización sería mostrar que la restricción de generalidad tiene contraejemplos.<sup>14</sup> Si esto es así y si se suponía que la restricción explicitaba (parte de) lo que está involucrado en la noción de contenido conceptual, entonces nos quedaríamos con una noción de contenido conceptual vacía. Y en tanto que la noción de contenido no-conceptual es derivada de la de contenido conceptual, no habría distinción fundada entre ambas

---

que es de la forma 'si S entonces C', y que hay un mecanismo causal que lleva al individuo a C cuando está en S/C y S.

<sup>13</sup> Podría darse el caso de que los ESs de un sistema de procesamiento de información cumplieran la restricción de generalidad, pero su cumplimiento sería contingente, dado que tal como se dice más arriba y a continuación, no es una condición esencial del contenido de los ESs la posesión de conceptos.

<sup>14</sup> Richard (1994, pp. 310-15) presenta un contraejemplo a la restricción y también objeta la versión de la misma de Davies (1989).

nociones. Esta estrategia no está nada mal. Lo que podría hacer un defensor de la distinción es ofrecer un enfoque para dar cuenta de en qué consiste tener los recursos (capacidades) conceptuales para tener un pensamiento particular  $Fa$  (o como suele desplegarse esto: en qué consiste que un individuo está en condiciones de conocer lo que hace que cualquier objeto arbitrario sea  $Fo$ , lo que es lo mismo, que cierto pensamiento  $Fa$  sea verdadero) que por supuesto no esté ligado a (no suponga) la restricción de generalidad. Y esto parece difícil de lograr.<sup>15</sup>

Más allá de las dificultades que pueda haber para dar cuerpo a la noción de contenido conceptual,<sup>16</sup> es innegable que la distinción conceptual- no-conceptual parece recoger una intuición muy fuerte acerca de las capacidades cognitivas humanas.<sup>17</sup> La intuición afirma que la especie humana comparte ciertas capacidades cognitivas (inferiores) con otras especies, y algunas (las superiores) son sólo propias de los humanos. Si hay algo que nos distingue claramente del resto de los animales son nuestras capacidades lingüísticas y dado que en general las teorías acerca del pensamiento han sido moldeadas en función de ciertas propiedades de nuestro lenguaje, no resulta extraño pensar que así como dominar un lenguaje requiere dominar términos, y dado que el lenguaje expresa los pensamientos, los conceptos (que son la contrapartida mental de los términos) son los constituyentes del pensamiento y en este sentido todo contenido intencional es contenido conceptual.<sup>18</sup> Por otro lado, si hay algo que nos equipara con el resto de las especies es la capacidad de procesar información. Una de las motivaciones para la postulación de contenido no-conceptual es la idea de que hay estados de

<sup>15</sup> Cf. la objeción de Richard (1994) al intento de Peacocke (1992).

<sup>16</sup> Por supuesto que en la literatura sobre el tema hay varias definiciones de contenido conceptual (y no-conceptual), la cuestión es que esas definiciones requieren de un enfoque acerca de la posesión de conceptos y es justamente este enfoque el que muy probablemente apele a una u otra variante de la restricción de generalidad.

<sup>17</sup> También está asociada a una distinción epistemológica muy arraigada en toda la historia de la filosofía y que ha recibido varias etiquetas: intuiciones-conceptos, datos de los sentidos-conceptos, esquema conceptual-contenido, etc. La conexión entre la distinción psicológica conceptual-no-conceptual y las distintas maneras de entender la tradicional distinción epistemológica se ve más claramente en la cuestión del contenido de la experiencia perceptual. Véase, por ejemplo, McDowell (1994a).

<sup>18</sup> Esto sería así para un enfoque de los conceptos como particulares mentales o habilidades con una base categorial (mecanismo causal) subyacente.

procesamiento de la información que compartimos con otras especies y para los cuales estar en ellos no requiere que el organismo domine conceptos.<sup>19</sup>

Supongamos que quisieramos respetar la intuición fuerte de una distinción entre contenido conceptual y contenido no-conceptual (debido a sus legítimas motivaciones evolutivas, ontogenéticas, y demás), y que aceptamos que en el dominio intencional, el contenido de los estados mentales está conceptualizado. Aun así, la característica de poseer contenido no-conceptual no clasifica a un estado como subdoxástico. Si la distinción entre el contenido conceptual y el contenido no-conceptual fuera realmente una distinción de principio entre ESs y estados intencionales, la sola presencia de contenido conceptual haría (sería una condición suficiente para) que un estado clasificara en el ámbito intencional y la sola presencia de contenido no-conceptual haría que un cierto estado clasificara como subdoxástico. Sin embargo, si bien por mor del argumento la primera condición parece cumplirse, la segunda no.

Hay quienes consideran que las experiencias perceptuales tienen contenido no-conceptual, entre ellos, el propio Davies (1991, 1995). Dado que el contenido no-conceptual sería la marca de lo subdoxástico, las experiencias perceptuales serían ESs. Sin embargo, esto no tiene sentido. Los ejemplos de ESs perceptuales que figuran en la literatura sobre la visión, que seguramente constituye la capacidad perceptual más y mejor estudiada, se refieren a los estados perceptuales del procesamiento visual temprano, que no constituyen por sí mismos experiencias perceptuales. Nadie, hasta donde sé, considera a las experiencias como meramente procesamiento de información, todos, hasta donde sé, distinguen entre ESs perceptuales y experiencias perceptuales.<sup>20</sup>

---

<sup>19</sup> Hay otras motivaciones relacionadas con las experiencias perceptuales. En particular, la idea de que el contenido experiencial es de grano más fino que los conceptos perceptuales (podemos discriminar tonos de colores para los cuales podríamos no tener conceptos de color) y que el contenido experiencial es independiente de las capacidades conceptuales (en el caso de las ilusiones visuales, éstas persisten aún no creyendo lo que se ve). Véase McDowell (1994a, cap. 3) para una discusión de las motivaciones mencionadas. También hay motivaciones relacionadas con elementos demostrativos e indexicales contenidos en los estados representacionales, de manera explícita o implícita. Véase Perry (1979), Peacocke (1986) y Cussins (1990).

<sup>20</sup> Esto se ve explícitamente en Bermúdez (1995) y Evans (1982), §6.3 y §7.4. Asimismo, véase la nota 8. Con "experiencia perceptual" me refiero a cómo las cosas le parecen al individuo o a percibir algo como así y asá, son estados conscientes de un individuo para los cuales además de los estados de

Por otro lado, las experiencias perceptuales tienen características que no poseen los ESs, en particular, están ligadas a la conciencia fenoménica. Tienen un componente fenoménico, que consiste en cómo es estar en esa experiencia, que no acompaña a ningún estado de procesamiento subdoxástico de la información. No hay algo así como qué se siente estar en el ES que consiste en un cierto arreglo de rasgos lexicales (ni tendría sentido que lo hubiera, si es que hay algo como qué se siente estar en un estado en particular). En verdad, el rasgo fundamental que distingue a estados con contenido no-conceptual como los subdoxásticos y las experiencias perceptuales, que supuestamente tendrían contenido no-conceptual es que los primeros están en el nivel subpersonal mientras que los segundos están en el nivel personal. Es ampliamente aceptado que los estados intencionales y las experiencias perceptuales son estados adscribibles a la persona y no a ninguna parte de la persona. Son las personas, en tanto sistemas cognitivos globales quienes creen, desean, y similares, y quienes experimentan dolor y demás. De manera que el defensor del criterio (c), debería apelar a un criterio ulterior basado en la distinción personal-subpersonal.<sup>21</sup>

La discusión de Davies (1989) de los criterios (a)-(c) iba dirigida a explorar las diferencias intuitivas entre los ESs y los estados intencionales de manera de ver si se podía hacer un caso para una distinción fundada. Si uno hace una lectura fuerte de Davies (1989) el objetivo sería concluir que es el criterio de conceptualización el que carga todo el peso de la distinción.<sup>22</sup> Pero, si la búsqueda de un criterio que pueda proveer una distinción de principio quiere decir la búsqueda de una condición necesaria y suficiente para que algo sea (o clasifique) como un ES, el contenido-no conceptual no califica puesto que, tal como hemos visto, no es una condición suficiente. Quizá sea una condición necesaria, pero no voy a explorar aquí esta opción.<sup>23</sup> Si uno hace una lectura más débil, la idea es que

---

información (ESs) se requiere de la interacción de los mismos con otros subsistemas como la memoria (o experiencia previa) y la atención.

<sup>21</sup> Para la distinción personal-subpersonal, véase el capítulo 2 (§1.1 y §1.2) y la conclusión.

<sup>22</sup> De hecho en Davies (1989, p. 148; 1995, p. 376) se dice que es este criterio el que distingue entre los ESs y los estados intencionales.

<sup>23</sup> Si se entiende contenido-no conceptual en el sentido de que no hace falta que el individuo posea los conceptos con los que se describe cierto contenido de un estado cognitivo para estar en él, dado que los conceptos no son constitutivos del estado, está claro que parece ser una condición

el criterio de conceptualización suma a los dos restantes en pos de una distinción fundada.<sup>24</sup>

Más allá de que los criterios (a)-(c) sean los adecuados para una distinción entre estados subdoxásticos y doxásticos, hemos llegado a la conclusión de que se requiere además de una distinción entre aquellos estados que forman parte de subsistemas de la persona y aquellos que forman parte de la persona en su totalidad.<sup>25</sup> Sin embargo, con esta distinción sola no alcanza para una diferenciación entre los estados intencionales y los no-intencionales. Hay quienes sostienen que incluso los estados subpersonales son intencionales en el sentido de que su especificación requiere apelar a propiedades distales de los objetos en el mundo exterior al individuo.

En el §2, voy a introducir otro criterio que permitiría, en principio, distinguir los estados intencionales personales de los subpersonales no-intencionales. Este criterio establece que los estados subpersonales tienen un 'contenido interno' a diferencia del 'contenido intencional' de los estados intencionales.<sup>26</sup> Mi propuesta puede ser leída de la misma manera que la de Davies. En un sentido fuerte, es el criterio de 'contenido interno' el que carga todo el peso de la distinción entre los estados intencionales personales y los no-intencionales subpersonales o bien, en un

---

necesaria de los ESs lingüísticos (cf. Chomsky 1986, p. 267). Sin embargo, cuando se trata de dar un enfoque sustantivo para especificar el contenido no-conceptual, se lo suele hacer apelando a capacidades o habilidades discriminatorias (por ejemplo, Cussins 1990). Así, se dice que los ESs de la visión involucran, en un sentido de discriminación, el reaccionar diferencialmente a algo en el campo visual. Más allá de que esto sea así, no queda nada claro en qué consistirían las habilidades discriminatorias de un ES que representa, entre otras propiedades, propiedades sintácticas de las expresiones lingüísticas. (Hay otras formas de especificación del contenido no-conceptual, por ejemplo, la de Peacocke 1992, pero en este caso aplicada a un determinado tipo de contenido: el espacial. Es mucho menos claro cómo se podría extender esta especificación al contenido no-conceptual de otros sistemas cognitivos subdoxásticos.)

<sup>24</sup> La intención de Davies (1989) parece ser la fuerte. Los pasajes que abren y cierran el análisis del criterio (c) avalarían esta lectura: "La propuesta a considerar ahora es que la restricción de generalidad ... puede proveer una distinción de principio entre el dominio de las actitudes proposicionales y el dominio subdoxástico" (p. 147) y "Lo que surge de la propuesta es que una distinción de principio entre estados de actitud y estados subdoxásticos tiene que fundarse en una distinción entre dos nociones de contenido" (p. 148).

<sup>25</sup> En lo que sigue hablaré indistintamente de estados y mecanismos subdoxásticos o doxásticos y subpersonales o personales, asimilando la primera distinción a la segunda.

<sup>26</sup> Hay que entender 'contenido intencional' en el sentido especificado en el capítulo anterior. Esto es, el contenido de los estados mentales especificado en el lenguaje natural por una oración de AP (vs. representaciones con propiedades intencionales).

sentido débil, este criterio se suma a los otros a favor de una distinción mejor fundamentada.

## 2. Contenido interno y contenido intencional

Hoy en día, se considera que los estados lingüísticos son subdoxásticos al igual que los de la visión (Davies 1986, 1989; Manfredi 1993; Bermúdez 1995; Egan 1995). Hay un sentido en el que el mecanismo computacional del lenguaje procesa y almacena información acerca de las propiedades semánticas, sintácticas y fonéticas de las expresiones lingüísticas que es utilizada por los subsistemas de ejecución de manera de dar lugar al uso del lenguaje (Chomsky 1995a, p. 2) y a creencias acerca del lenguaje (Chomsky 1980b, p. 57). Veamos en qué consisten el mecanismo y los estados lingüísticos en la teoría chomskiana, para luego, a partir de ciertas consideraciones chomskianas, introducir la noción de 'contenido interno' para los estados subpersonales.

La (psico)lingüística generativa centra el estudio del lenguaje en el funcionamiento de un subsistema de la mente humana denominado facultad del lenguaje. La facultad del lenguaje tiene dos componentes: un sistema cognitivo (que almacena información) y sistemas de ejecución (perceptivo-articulatorio e intencional-conceptual). La (psico)lingüística se ocupa del estado inicial y el estado estable del sistema cognitivo. La 'gramática universal' es el estudio acerca del estado inicial y la 'gramática' se ocupa del estudio del lenguaje-I que consiste en el estado estable alcanzado por los hablantes, esto es, su competencia lingüística. La gramática universal es la teoría acerca de los lenguajes-I humanos posibles. Postula un conjunto finito de principios, que son parte del equipamiento biológico humano, con parámetros cuyos valores se fijan en función de la experiencia lingüística durante la adquisición del lenguaje. Las lenguas particulares (los lenguajes-I) son distintas instanciaciones según las distintas parametrizaciones de los principios de la gramática universal.

La noción técnica de lenguaje-I supone una cierta idealización que consiste en hablantes con experiencia uniforme en comunidades lingüísticas homogéneas (sin diversidad de dialectos ni variaciones entre hablantes). Fue acuñada para diferenciar el objeto de la lingüística de los desarrollos anteriores (la lingüística descriptiva y estructural) que se focalizaban en el lenguaje-E. El lenguaje-I es Interno (el objeto de estudio es interno a la mente/cerebro), Individual (de hablantes particulares) e Intensional (caracterización en intensión de la función que genera descripciones estructurales). En cambio, el lenguaje-E es independiente de las propiedades de la mente/cerebro (externo), consiste en una enumeración de descripciones estructurales (extensión) y no es individual en el sentido de que no se relaciona con estados mentales de individuos.<sup>27</sup> El lenguaje-I (su estado inicial y su estado estable) constituye un objeto real en las mentes/cerebros de los hablantes, el lingüista estudia aquellos aspectos abstractos del cerebro, esto es, los mentales (lingüísticos). Pero, en tanto su objeto de estudio es un objeto natural, la lingüística forma parte de las ciencias naturales y encara su objeto con la misma metodología que emplean el resto de las ciencias naturales.

Según el programa minimalista (Chomsky 1995a, 1995b), cada lenguaje-I consiste en un mecanismo computacional y un léxico. El procedimiento computacional toma como inputs instancias de eventos físicos (las ondas sonoras de las expresiones lingüísticas), computa sobre rasgos lexicales y sus outputs son objetos abstractos (simbólicos o formales), que llevan en sí toda la información semántica, sintáctica y fonética necesaria para asociar el sonido y el significado de las expresiones lingüísticas. Es decir, que genera descripciones estructurales de las expresiones lingüísticas a partir de elecciones léxicas. Las descripciones estructurales consisten en un complejo de propiedades fonéticas, sintácticas y semánticas que actúan como instrucciones para los sistemas de ejecución: el articulatorio-auditivo y el conceptual-intencional. El léxico es una lista de ítems léxicos idiosincráticos (que no se derivan de principios generales). Los ítems léxicos están caracterizados por una serie de rasgos semánticos, sintácticos y fonéticos. Por

---

<sup>27</sup> Para la distinción entre lenguaje-I y lenguaje-E, véase, entre otros, Chomsky (1986), cap. 2; (1991) y Smith (1999, pp. 138-9).

ejemplo, la entrada lexical de 'libro' tiene, entre otros, el rasgo semántico [artefacto], el rasgo formal [sustantivo] y el rasgo fonético [comienza con consonante oclusiva].<sup>28</sup>

El procedimiento computacional toma ítems léxicos cuyos rasgos fonéticos se computan para dar lugar a una representación fonética (o FF), toma rasgos semánticos para dar lugar a una representación semántica (o FL) y los rasgos formales entran en la computación pero deben ser eliminados (en el camino a la FF) para dar lugar a una representación convergente, esto es, interpretable por los sistemas de ejecución de la interface. Son estos sistemas de interface los que imponen condiciones de legibilidad puesto que son los que interpretan el output y siguen sus instrucciones para el pensamiento y la acción. El mecanismo computacional toma ítems léxicos y computa sobre sus rasgos hasta un punto (denominado *Spell-Out*) en donde las derivaciones se separan hacia los dos niveles de interfaz. Las operaciones que se realizan entre *Spell-Out* hasta concluir en la FF no son de la misma naturaleza que las que llevan a la FL, se asume que el mecanismo computacional lleva a la FL mientras que un componente separado, el componente fonológico, es el que realiza las operaciones hasta la FF.

También se asume que no hay interacciones entre la FL y la FF, y que el mecanismo computacional es derivacional, esto significa que hay operaciones sucesivas que llevan a la FL y FF. En este sentido, los únicos niveles de representación son la FL y FF. Las operaciones sucesivas del mecanismo lo que hacen es chequear rasgos (de caso, tiempo, concordancia), las estructuras que no cumplen con las condiciones de la interface no convergen (y no pueden ser interpretadas por los sistemas de ejecución). Del conjunto de derivaciones convergentes (que satisfacen las condiciones de interfaz) hay un subconjunto de derivaciones admisibles seleccionadas por principios de economía del sistema.<sup>29</sup>

---

<sup>28</sup> Chomsky acepta la hipótesis que afirma que los principios parametrizados de la gramática universal no se relacionan con el mecanismo computacional sino con el léxico, en ese sentido cada parámetro está relacionado con elementos del léxico de manera que las variaciones entre los lenguajes estarían dadas por parámetros lexicales, en particular, morfológicos (Chomsky 1995a, p. 131).

<sup>29</sup> Los principios de economía que se aplican en cualquier punto de la derivación son: *shortest move*, *greedy* y *procrastinate*. Ver Chomsky (1995a).



De manera que, una expresión lingüística E consiste, al menos, de un par {FON, SEM} formado por una derivación convergente óptima. FON(E) es una representación de la FF que contiene información relevante al sonido de E y SEM(E) es una representación de la FL que porta información relevante al significado de E. Se asume que FON(E) y SEM(E) son objetos legítimos (el tipo de representación legítima de la FF es un problema estándar de la fonética universal, el tipo de representación legítima de la FL se supone que es una cadena  $(CH = (\alpha_1, \dots, \alpha_n))$  que al menos tiene un núcleo, un argumento, y un modificador) que reciben interpretación por los sistemas perceptual-articulatorio y el intencional-conceptual, respectivamente.

Que un hablante conozca un cierto lenguaje-I consiste en tener un mecanismo cognitivo de esta naturaleza. Todo hablante de cualquier lengua tiene una estructura mental que funciona de manera de proporcionarle la información adecuada para dominar su lengua. Es en este sentido que el lenguaje-I constituye un sistema de conocimiento (competencia lingüística).<sup>30</sup> Ahora bien, ha habido mucho debate en torno a en qué sentido de “conocer” un hablante conoce hechos lingüísticos, al punto que Chomsky (1975) tuvo que introducir el neologismo

---

<sup>30</sup> Los hablantes conocen hechos acerca de (a) los principios de la gramática universal, (b) la gramática de su propia lengua (derivados de la parametrización de (a)) y (c) propiedades de las emisiones (derivadas de (a) y (b)) (Chomsky 1986, p. 265; 1980a, p. 80). Un ejemplo de (a) consiste en que dada dos oraciones (1) *He thinks Bill is a nice guy* (El piensa que Bill es un muchacho agradable) y (2) *The woman he married thinks Bill is a nice guy* (La mujer que se casó con él piensa que Bill es un muchacho agradable), un hablante conoce que en (1) *él* no depende referencialmente de *Bill* (no se interpreta (1) como (1') Bill piensa que Bill es un muchacho agradable), mientras que en (2) es posible la dependencia referencial (se puede interpretar (2) como (2') La mujer que se casó con Bill piensa que Bill es un muchacho agradable). Esto se explica porque el estado inicial de la facultad del lenguaje tiene principios relacionados con la dependencia referencial (que son estudiados por la Teoría del Ligamiento. Chomsky 1994a, p. 93).

Un ejemplo de (b) consiste en que un hablante del español conoce que cuando el objeto del verbo es animado va precedido de la preposición 'a' (como en 'Juan afeita a Pedro') sin embargo conoce que no puede realizar construcciones como 'Juan hizo afeitar a Pedro a María' (a diferencia de 'Juan hizo arreglar el carro a María'). Esto se debe a que el español (pero no el italiano, por ejemplo) no permite que haya dos sintagmas nominales precedidos de 'a' (ambas 'a' no constituyen sintagmas preposicionales verdaderos sino que están por razones sintácticas, mientras que en 'Juan tiró a su amigo al agua' está permitido porque uno de los sintagmas precedidos por 'a' es un sintagma preposicional verdadero. Chomsky 1988, pp. 20-1).

Un ejemplo de (c) sería que un hablante del inglés conoce que “en la palabra inglesa *pin*, /p/ es aspirada mientras que en *spin* no lo es” (Chomsky 1986, p. 266), que *house* rima con *mouse* y que '*brown house*' se usa para referir a una estructura diseñada con cierto propósito con una superficie marrón (Chomsky 1995b, p. 14).

“*cognize*” para aludir a un tipo de conocimiento proposicional inconciente, tácito que no posee parte de las propiedades que la tradición epistemológica le atribuye (por ejemplo, no requiere justificación ni verdad ni es conciente).<sup>31</sup>

No voy a entrar en esta polémica porque creo que todo lo que se dice es que la competencia de un hablante consiste en tener un lenguaje-I que es tener un procedimiento generativo (en la mente/cerebro) que asigna descripciones estructurales a expresiones lingüísticas. Así, tener un lenguaje consiste en la explicación de cómo funciona el mecanismo del lenguaje-I. Más allá de cualquier aspecto controvertido acerca de la cuestión del conocimiento, la explicación del funcionamiento del mecanismo se mantiene en pie, y los estados del lenguaje-I son estados que almacenan información lingüística a la que no se tiene acceso conciente.

Si tenemos que hablar de los estados del mecanismo computacional del lenguaje-I, podríamos hacer un corte en cualquier punto de las derivaciones y considerar esos estados intermedios o bien el output del mecanismo, o ambos tipos de estados. Por cuestiones de simplificación y dado que los únicos niveles de representación admitidos por la teoría minimalista son la FL y la FF, lo atinado sería tomar en cuenta estos estados del output, que, en última instancia, consisten en una disposición de rasgos lexicales (Chomsky 1995a, p. 225). Me referiré a ellos como estados lingüísticos (esto no significa que no se puedan considerar estados

---

<sup>31</sup> Por un lado, dentro del proyecto naturalista chomskiano no hay ninguna relación epistémica sustantiva entre el individuo y el lenguaje-I, tal como la epistemología tradicional entiende la relación de conocimiento. En este sentido acuerdo con Stone y Davies cuando sostienen que las restricciones filosóficas que impone la epistemología son ociosas para un proyecto naturalista como el chomskiano en el que:

No se ofrece ninguna respuesta a cuestiones acerca de la naturaleza epistemológica de la relación entre una persona y un lenguaje-I. En teorías naturalistas acerca del lenguaje no hay una cuestión así para ser respondida, así como no hay una cuestión a responder en la teoría de la percepción visual acerca de la relación epistemológica entre una persona y los principios que capturan el funcionamiento del sistema visual. (2002, p. 281)

Por otro lado, esta noción es bastante utilizada en psicología cognitiva en el sentido de “información que es tratada por el sistema que la usa como si fuera fidedigna” (Cummins 1997, p. 536).

intermedios, sólo significa que los estados lingüísticos propiamente dichos serán los del output del mecanismo del lenguaje-I).

Ahora bien, según Chomsky, el estudio del lenguaje-I es naturalista e internalista. Lo primero apunta al “estudio de los humanos como lo hacemos con cualquier otra cosa en el mundo natural” y lo segundo a “una investigación que busca comprender los estados internos de un organismo” (Chomsky 1995a, p. 27). Así, el estudio del lenguaje-I es acerca del mecanismo computacional, las representaciones mentales sobre las que opera y las relaciones entre ellas. Chomsky suele referirse a este tipo de estudio como sintaxis en sentido amplio, esto es, el estudio de las representaciones mentales y las computaciones *vs.* un estudio acerca de las relaciones entre las representaciones y el mundo (Chomsky 1995a, p. 19; 1991, p. 38) o como un estudio en el nivel simbólico *vs.* en el nivel semántico-intencional (Chomsky 1986, p. 262; 1992a, p. 34).

Ambas expresiones, “sintaxis pura” y “nivel simbólico”, apuntan a que el estudio de los mecanismos cognitivos es completamente interno, en el sentido de que las descripciones que se utilizan para describir los estados internos (y el contenido de esos estados) y las funciones que se computan no hacen referencia a ningún rasgo del mundo externo a la persona.<sup>32</sup> El ámbito de este tipo de investigación se restringe a las transformaciones por las que pasa un subsistema cognitivo desde su estado inicial (representaciones iniciales) a su estado estable (representaciones del output del subsistema). En este sentido, el estudio del mecanismo del lenguaje-I por el cual proyecta una serie de elecciones léxicas en un par de objetos simbólicos, FF y FL, es pura sintaxis, al igual que “la investigación de cómo la imagen de un cubo rotando en el espacio está determinada a partir de la estimulación retinal o imaginada” (Chomsky 1995a, p. 19). De lo que se trata es del estudio de cómo funcionan los mecanismos cognitivos que están en la cabeza de los individuos, y de las propiedades y configuraciones de los objetos simbólicos que son las representaciones mentales, sin necesidad de establecer ningún tipo de conexión entre éstas y el entorno.

---

<sup>32</sup> Este nivel simbólico es al que he aludido cuando he hablado del nivel de descripción semántico como diferente del intencional. Volveré sobre esto más adelante.

Sin embargo, la mayoría de los filósofos consideran que los estados mentales de los mecanismos subdoxásticos poseen un contenido semántico a la manera de los estados intencionales, sin ser ellos mismos estados intencionales de AP (por las razones vistas en §1). La creencia que  $p$  y el ES que representa (o almacena) la información que  $p$  tienen contenido semántico (portan información acerca del mundo) y en ese sentido, son intencionales (Davies 1989, Peacocke 1994). A veces se suele hablar de “acerquidad” (“*aboutness*”, Davies 1995; Bermúdez 1995) para marcar cierta diferencia con los estados de AP que serían intencionales de pura cepa.<sup>33</sup> Hay quienes específicamente sostienen que las representaciones postuladas por Marr en su teoría de la visión son acerca de los objetos que las causan y en ese sentido consideran que el contenido de las representaciones está determinado por cómo es el mundo, esto es, sostienen una posición externalista con respecto a los contenidos de los ESs (Burge 1986, Davies 1991).

Ahora bien, si uno atiende a las teorías de los mecanismos subdoxásticos analizadas por estos filósofos, las teorías cuyos precursores son Chomsky y Marr, se puede advertir que se despliegan bajo las directivas de una investigación naturalista internalista. Me centraré en la teoría y los argumentos chomskianos y haré algunas observaciones acerca de la teoría de Marr.<sup>34</sup> La defensa chomskiana de un enfoque internalista se puede reconstruir a partir de tres argumentos principales.

En primer lugar, se puede distinguir un argumento general que intenta sostener que la noción de contenido intencional no forma parte de una teoría acerca de los sistemas cognitivos subpersonales. Dado que no es una noción teórica, la polémica filosófica internalismo-externalismo acerca de los contenidos

---

<sup>33</sup> Según Davies (1995), la intencionalidad de pura cepa o genuina de los estados de AP requiere de ciertos requisitos como la posesión de conceptos, el requisito de generalidad, y demás que hemos visto en §1. La intencionalidad de los ES, si bien no requiere de esas condiciones, no es intencionalidad ‘como si’ porque los ESs tienen propiedades intencionales genuinas. Stich (1978) dice que la creencia de que  $p$  y el ES que representa la información de que  $p$  se parecen en las causas (también Evans 1981) o en alguna parte del patrón inferencial (Evans 1981 dice que tienen consecuencias inferenciales cuyos outputs son creencias) o en algún efecto no inferencial que puedan tener o en alguna combinación de estos aspectos.

<sup>34</sup> Aunque sin pretensiones exhaustivas puesto que no es este el lugar para desarrollar una defensa detallada del internalismo de la teoría de Marr. Quienes sí la han desarrollado son Egan (1992, 1999), Segal (1989, 1991) y McGinn (1989, pp. 58-100).

mentales es estéril cuando se intenta aplicarla a los sistemas subdoxásticos. Segundo, hay un argumento específico que trata de mostrar lo mismo que el primero, pero limitado sólo al subsistema cognitivo lingüístico. Se basa, al menos, en dos ideas específicas: la idea de que una expresión no existe aparte de sus propiedades en el nivel de interfaz (Chomsky 2000, p. 175) y el requisito de mantener cierto paralelismo en la explicación de las propiedades semánticas y fonológicas de las expresiones lingüísticas (Chomsky 2000, pp. 175-80). Y, tercero, hay un argumento que sostiene que las cuestiones sobre intencionalidad quedan fuera de una investigación naturalista internalista. Este argumento se basa en las ideas de que el fenómeno de la intencionalidad no puede ser objeto de estudio de ninguna disciplina empírica (Chomsky 1995a, p. 27; 1992a, p. 22) y que quizá su estudio constituya un misterio (Chomsky 1995a, p. 27).

Las relaciones entre los argumentos son las siguientes. El primer argumento implica el segundo pero no el tercero. Esto es, acordar con la idea de que la noción de contenido intencional no juega ningún rol en las teorías subpersonales implica suscribir al argumento específico acerca del subsistema lingüístico pero no a la idea de que la intencionalidad no puede ser objeto de estudio (al menos parcial) de las disciplinas empíricas (y por ende, no implica adherir a la idea de que su estudio constituye un misterio). El segundo argumento no implica ni el primero ni el tercero. Se podría aceptar que en el caso de la teoría lingüística no hay una noción de contenido intencional, pero que en otras teorías acerca de otras capacidades podría haberla, y por ende, no es cierto que la intencionalidad cae fuera de un estudio empírico. Por otro lado, aceptar el tercer argumento implica la aceptación de los demás. Si la intencionalidad no puede ser el objeto de estudio de las disciplinas empíricas, se sigue que la noción de contenido intencional no forma parte ni de una teoría acerca del subsistema lingüístico ni de los subsistemas cognitivos en general.

Como me interesa abordar la cuestión general acerca de si se requiere una noción de contenido intencional en las teorías subpersonales, en lo que sigue no voy a abordar el argumento específico acerca de la capacidad lingüística sino que me

concentraré en el primero y el segundo.<sup>35</sup> Mi intención va a ser defender la idea de que la noción de contenido intencional no forma parte de las teorías subpersonales, sin adscribir a la idea de que la intencionalidad no puede ser objeto (parcial) de estudio de las disciplinas empíricas. En esta sección me ocuparé del primer argumento y en el §4 me ocuparé del segundo y haré algunas observaciones en torno al dominio psicológico.

Hemos visto que el output del lenguaje-I para una expresión lingüística E consiste en el par {FON, SEM}. FON(E) y SEM(E) son representaciones cuya información consiste en propiedades fonéticas y semánticas, respectivamente, de E. Estas representaciones fonética y semántica no representan nada fuera de los sistemas que hacen uso de ellas, no son acerca de nada en el mundo exterior, sino que son propiedades internamente determinadas por el lenguaje-I (más exactamente, por el estado inicial de la facultad del lenguaje) que sirven de instrucciones para los sistemas de ejecución. Estas representaciones llevan información completamente interna al mecanismo del lenguaje-I y en este sentido, no hay algo así como un contenido fonético determinado (e individuado) por cómo es el mundo, ya sea por “movimientos moleculares o convenciones acerca de la “pronunciación apropiada”” (Chomsky 1995a, p. 44). Análogamente, no hay un contenido de las representaciones semánticas que esté determinado por el mundo. En palabras de Chomsky:

El estudio internalista del lenguaje también habla de “representaciones” de varias clases, incluyendo representaciones fonéticas y semánticas en la “interfaz” con

---

<sup>35</sup> El argumento específico alude a la distinción entre el lenguaje-I y el lenguaje-E que se ha visto más arriba. Dado que el único objeto real es el lenguaje-I, la tarea de conectar las representaciones que intervienen en el mecanismo computacional del lenguaje-I con algo del mundo no tiene razón de ser, no habiendo nada en el mundo con que conectarlas. Además, la idea de que en el caso de ciertas representaciones tiene sentido hablar de contenido intencional pierde su fuerza cuando se toma en cuenta los aspectos fonológicos del output del lenguaje-I. Chomsky (2000, pp. 176-80) considera hipótesis alternativas basadas en el mentalés (correspondencia de los elementos léxicos con elementos del mentalés), teorías denotacionales (los elementos léxicos denotan un valor semántico externo al individuo), teorías en términos de creencias (las propiedades de una expresión se explican por creencias del hablante acerca del mundo), y teorías de doble factor (las propiedades de una expresión se explican en términos del rol inferencial más una teoría denotacional). Con cualquiera de estas hipótesis alternativas se pierde el paralelo en la interpretación de las representaciones semántica y fonológica (sobre todo porque no hay propuestas de la mayoría de ellas para el aspecto fonológico).

otros sistemas. Pero aquí tampoco necesitamos reflexionar acerca de lo que está representado, buscando alguna construcción objetiva de sonidos o cosas. Las representaciones son entidades mentales postuladas que tienen que entenderse a la manera de una imagen mental de un cubo rotando, ya sea que sea consecuencia de presentaciones taquistocópicas o de un cubo real rotando o de la estimulación de la retina por algún otro medio; o, incluso, de la imaginación. Las representaciones internas del lenguaje, en tanto accedidas por los sistemas de ejecución, participan en la interpretación, el pensamiento y la acción, pero no hay razón para buscar *ninguna otra relación con el mundo*, como podrían sugerir una conocida tradición filosófica y analogías inapropiadas del uso informal. (1995a, p. 53, mis itálicas)

La cita alude a, al menos, dos cosas. En primer lugar, alude a la esencia de una investigación internalista, esto es, que se ocupa solamente de aquellos aspectos mentales (en este caso, los lingüísticos) que supervienen a estados y procesos internos al individuo, y en segundo lugar, hay una afirmación acerca de que las cuestiones sobre la intencionalidad surgen en un nivel distinto de análisis. Veamos lo primero. Las teorías que dan cuenta de los mecanismos lingüístico y visual se concentran en el estudio de las transformaciones entre las representaciones internas sin que sea pertinente el entorno en el que el individuo esté inmerso. Para lo que cuenta en una investigación internalista, podríamos ser “cerebros en la cubeta”, en este sentido sólo interesa aquello que superviene intrínsecamente al cerebro.

No hace falta apelar al entorno para describir la información que llevan las representaciones puesto que ellas portan información que sólo pueden leer los mecanismos pertinentes y cuyos outputs interpretan otros sistemas cognitivos del individuo. El estudio psicológico comienza con las representaciones iniciales y culmina con las representaciones que son el output para otros sistemas cognitivos. Así, el mecanismo del lenguaje-I proyecta una serie de elecciones léxicas en un par de objetos simbólicos: FF y FL, y el mecanismo de la visión proyecta imágenes

retinales en la corteza visual.<sup>36</sup> En este sentido, no hay ninguna noción interesante de contenido relacionada con los mecanismos subpersonales, no al menos si se la entiende en términos intencionales.

En segundo lugar, el subrayado en la cita da a entender que si hay una conexión con el mundo, ésta no se da en el nivel subpersonal sino en el nivel personal del pensamiento y la acción. Los estados lingüísticos que son el producto del lenguaje-I son el output de un subsistema muy específico de la arquitectura mental. Estas representaciones tienen que ser interpretadas por los sistemas de ejecución lingüística: el perceptivo-articulatorio y el intencional-conceptual. La interacción con estos sistemas de ejecución, que a su vez se relacionan con los intereses, propósitos y conocimiento general de los hablantes, permite que las representaciones lingüísticas así enriquecidas y transformadas participen de la comunicación lingüística, el pensamiento y la acción.

Tal como hemos visto en el capítulo anterior, el uso del lenguaje está relacionado con las personas (no con subsistemas aislados de las personas) y sus perspectivas, esto es, con sus intereses, propósitos, sus puntos de vista, todas variables sumamente complejas y que no pueden abordarse desde una concepción internalista. Es aquí donde el fenómeno de la intencionalidad tiene su lugar natural y consecuentemente es aquí, en el nivel personal, donde es atinente buscar la conexión con el mundo, y en este sentido, donde sería atinente hablar de contenido intencional. Pero, tal como se intentó mostrar también en el capítulo anterior, la fijación del contenido intencional no es una cuestión de conectar representaciones con aspectos del mundo sino que la conexión es mucho más compleja e intrincada.

---

<sup>36</sup> En el caso de la teoría de Marr la representación inicial que es el esbozo primario representa valores de intensidad en la retina, no se dice cómo se construye el esbozo primario (Cf. Segal, 1989, p. 194, n. 7). Por otro lado, se podría pensar que el output de la teoría de Marr (la representación del modelo de 3D) no constituye una representación que tenga que ser interpretada por, o que tenga que entrar en interacción con, otros mecanismos cognitivos (en otras palabras, se podría pensar que el output constituye de por sí la percepción del mundo objetivo). Sin embargo, la teoría de Marr (que sólo aborda un aspecto de la visión, las propiedades espaciales) es justamente una teoría en el nivel subpersonal porque sus outputs no están en el nivel personal. Esto es así, porque hace falta más para la experiencia perceptual personal (de ver algo como esto y aquello) que la teoría de Marr (se requiere conexión con sistemas como la memoria, la atención, con experiencias pasadas, y demás).



Se puede entender la idea de que las teorías internalistas no requieren una teoría del contenido intencional (y por ende las cuestiones sobre la intencionalidad no entran en una investigación naturalista) apelando a la distinción personal-subpersonal. En el capítulo 2 (§1.1.) se ha dicho que ésta es una distinción dentro del ámbito psicológico entre, por un lado, mecanismos y estados psicológicos que portan información significativa para la persona de manera que ésta puede usarla para el pensamiento y la acción, y por el otro, mecanismos y estados psicológicos que utilizan información que sólo está disponible para otros subsistemas de la persona. De acuerdo a esta distinción, lo que Chomsky estaría diciendo es que cualquier conexión que valga la pena de ser establecida con algo en el mundo (por más complicada que resulte esta tarea), no es en el nivel de los subsistemas subpersonales, sino en el nivel personal en donde entra en juego el ámbito de lo intencional. Con esto en mente, quedan más claros párrafos como el siguiente:

*No hay una cuestión que tenga sentido acerca del "contenido" de las representaciones internas de una persona viendo un cubo bajo condiciones experimentales, o si la retina está estimulada por un cubo rotando, o por un video de un cubo rotando; o en los experimentos estándar de la visión de las ranas, acerca del contenido de la "representación de" una mosca o de un punto que se mueve. Ninguna noción como "contenido", o "representación de" figura en la teoría, por eso no hay respuestas para dar acerca de su naturaleza. Lo mismo es verdad cuando Marr escribe que está estudiando la visión como "una proyección desde una representación a otra, y en el caso de la visión humana, la representación inicial consiste en, sin ninguna duda, series de valores de intensidad de la imagen tal como la detectan los fotorreceptores en la retina" (Marr, 1982, p. 31), donde "representación" no tiene que entenderse relacionamente como "representación de". (Chomsky 1995a, pp. 52-3, mis itálicas)*

Hay dos maneras de entender la idea de que en las teorías subpersonales no hay una noción de contenido intencional. O bien no la hay porque los científicos se manejan muy bien en sus investigaciones sin necesidad de hablar de contenido pero convendría para una comprensión teórica completa una noción así aportada

quizá por la filosofía de la psicología, o bien no la hay porque no es pertinente, esto es aún cuando hubiera una noción satisfactoria de contenido intencional no cumpliría ningún papel en la teoría. Creo que Chomsky está diciendo lo segundo.<sup>37</sup> La noción de contenido intencional no tiene un lugar en una teoría subpersonal porque de lo que se trata es de dar cuenta de cómo funcionan los mecanismos cognitivos y para ello no hace falta ninguna teoría acerca de si las representaciones del organismo son verdaderas o falsas, los mecanismos funcionan de la misma manera. Explicar cómo un organismo representa verdaderamente su entorno, o el éxito de las interacciones del organismo con su entorno depende de cuestiones que van más allá de una investigación psicológica de mecanismos subpersonales.

Ahora bien, la propuesta que surge de esta interpretación de las reflexiones y la teoría chomskianas es que los estados subdoxásticos subpersonales poseen un 'contenido interno'.<sup>38</sup> Este es el rasgo que permite establecer una distinción de principio entre los estados subpersonales y los estados intencionales personales. El contenido interno alude a la información interna que llevan ciertas representaciones, información de la que hacen uso o pueden decodificar sólo otros subsistemas cognitivos.<sup>39</sup> Las representaciones de los estados subpersonales portan información para los subsistemas, no para el sujeto. Es información que en el caso del lenguaje es accedida por los sistemas de ejecución y en el caso de la visión es accedida también por otros mecanismos. Los subsistemas que hacen uso de esta información interactúan entre sí de manera de dar lugar a estados intencionales y experiencias perceptuales que participan en el pensamiento y la acción. No tiene sentido buscar una conexión entre estas representaciones y el mundo, la conexión

---

<sup>37</sup> Egan (1992, p. 256) sostiene que sería erróneo atribuirle a Marr alguna teoría acerca del contenido. Esto serviría de apoyo para la idea de que la noción de contenido intencional no figura en una teoría subpersonal, sin embargo, Egan interpreta el contenido de los estados marrianos como contenido amplio. Discutiré esto en el §4.

<sup>38</sup> No estoy segura de que Chomsky acuerde en que la misma noción de 'contenido' (sin el adjetivo 'intencional') sea útil. Si esto es así, la propuesta va más allá y se aparta aquí de las consideraciones chomskianas.

<sup>39</sup> La noción de "información" que se utiliza aquí no es en el sentido de relaciones nómicas entre eventos tal como la entiende la semántica informacional de Dretske y Fodor. "Información" alude a las instrucciones que los componentes de un subsistema (o los subsistemas entre sí) intercambian de manera de ejecutar sus funciones.

se da, en todo caso, en los estados personales a que dan lugar, porque son aquellos cuyo contenido el sujeto puede manipular para el pensamiento y la acción.

La noción de contenido interno en tanto que sirve como un rasgo diferenciador de los estados subpersonales, tiene su lugar en la filosofía de la psicología, sin embargo a diferencia de otras nociones de contenido (por ejemplo, 'estrecho' y 'amplio'), no hace falta ninguna teoría sustancial acerca de la naturaleza del contenido interno. En tanto noción que participa de una teoría psicológica subpersonal, lo único que se requiere es que el investigador que propone la teoría describa la información interna que portan las representaciones postuladas. En este sentido, las representaciones tienen una interpretación, la que el teórico propone, pero ésta no es intencional porque no se alude a nada del mundo externo (propiedades distales u objetos).

Ahora bien, se podría pensar que la noción de contenido interno es una denominación distinta para lo que se ha llamado en la literatura 'contenido estrecho'. En el §3 mostraré que no es fructífero asimilar el contenido interno al estrecho en ningún sentido en que se ha desarrollado en detalle esta última noción.

### 3. Contenido interno y contenido estrecho

Si uno comienza relevando las distintas definiciones de contenido estrecho, habría algunas que podrían aplicarse perfectamente al contenido interno y en este sentido, se podría pensar que este último no es más que un ropaje nuevo para un viejo concepto. Así, es cierto que el contenido interno y el estrecho se asemejan en que supervienen a propiedades intrínsecas del individuo, y su individuación es independiente de la historia causal y del entorno del individuo. Sin embargo, si una parte de las nociones de contenido estrecho efectivamente desarrolladas se puede advertir que tanto las motivaciones para postular esta noción como el rol que cumple son completamente diferentes a la del contenido interno y en ese sentido, constituyen conceptos distintos. Lo que intentaré mostrar es que la noción de

contenido interno no es la misma que la de contenido estrecho. Esto es importante porque si uno asimila la noción de contenido interno a la de contenido estrecho, la primera no cumpliría con la tarea asignada de distinguir los estados subpersonales de los estados intencionales porque, fundamentalmente, la noción de contenido estrecho fue pensada para estados intencionales, en particular, para las creencias.

Por otro lado, si se deja a un lado el hecho de que la noción de contenido estrecho fue pensada para estados intencionales, de todas formas, aunque esta noción estuviera muy bien definida, no sería adecuada para aplicarla a los estados subpersonales. Esto es así porque la noción de contenido estrecho depende (en un sentido que se especificará más adelante) de la de contenido amplio y dado que los estados subpersonales no tienen contenido amplio, la noción de contenido estrecho pierde todo el sentido para la que fue pensada. De manera que la diferenciación entre el contenido estrecho y el interno se basa en dos tipos de consideraciones acerca del contenido estrecho. Por un lado, cuestiones generales acerca de las motivaciones para postularlo y el rol que cumpliría en las teorías psicológicas, y por otro lado, cuestiones relativas a sus relaciones con el contenido amplio. Comenzaré por las cuestiones generales.

La noción de contenido intencional fue pensada para los estados intencionales, en particular, las creencias y deseos. Su postulación estuvo ligada desde un comienzo a las explicaciones, de sentido común, de la conducta de los individuos. Las explicaciones intencionales atribuyen a los individuos estados intencionales cuyos contenidos tienen un rol causal en la producción de la conducta y son semánticamente evaluables. Experimentos mentales como los de la Tierra Gemela (Putnam 1975) mostraron que los contenidos de los estados mentales no podían cumplir al mismo tiempo con el requisito del rol causal y de la evaluabilidad semántica. Si es concebible que dos individuos molecularmente idénticos en entornos donde la única diferencia es la composición química del agua tengan conductas similares ante la misma, eso significa que el contenido que causa la

conducta del individuo no se ve afectado por las circunstancias que hacen que el mismo contenido sea verdadero o falso (esto es, por sus condiciones de verdad).<sup>40</sup>

De aquí surge la idea de la bifurcación del contenido, el contenido estrecho es el que interviene en las explicaciones psicológicas de la conducta mientras que el amplio está relacionado con motivaciones semánticas con respecto a la información acerca del mundo que portan las creencias o con el éxito de los individuos en sus interacciones con el entorno en virtud de que sus representaciones son verdaderas. La psicología taxonomiza los estados intencionales en términos de sus poderes causales y dado que estos supervienen sobre microestructuras locales (propiedades intrínsecas del organismo: físicas o sintácticas), sólo hay que apelar a propiedades intrínsecas para individuarlos (y, *a fortiori*, para individuar sus contenidos mentales). Aquello que comparten los gemelos en contextos distintos y lo que hace que tengan la misma conducta es el contenido estrecho. Por otro lado, dado que los estados portan información acerca del mundo, el contenido amplio es el que da cuenta de las propiedades relacionales del contenido. Algunos filósofos de la mente entendieron que la psicología necesitaba tanto del contenido estrecho como del amplio.

Ahora bien, quisiera enfatizar tres cuestiones. La primera es que la noción de contenido estrecho fue pensada para el nivel personal. Son los estados intencionales, que están en el nivel personal, los que causan la conducta en función de su contenido estrecho. En segundo lugar, el rol explicativo del contenido estrecho es dar cuenta de la conducta de los individuos. Y, en tercer lugar, si la conducta de los individuos es la misma, es porque comparten contenidos estrechos. O, dicho más específicamente, el contenido estrecho es aquello que en tanto que superviene sobre propiedades intrínsecas, es lo compartido por gemelos en entornos distintos. En cambio, la noción de contenido interno no se aplica a

---

<sup>40</sup> Además de los casos Putnam (identidad de contenido estrecho pero distintas condiciones de verdad), los casos Frege (idénticas condiciones de verdad pero distintos contenidos estrechos) también llevan a la bifurcación del contenido. El experimento mental de Putnam muestra, según él mismo y las distintas interpretaciones que tuvo, que el contenido no superviene a propiedades intrínsecas del individuo, esto es, muestra que el contenido es amplio. Sin embargo, muestra al mismo tiempo la necesidad de un contenido estrecho. Dado que los gemelos tienen conductas similares, la diferencia en la extensión de sus contenidos no afecta la conducta de los gemelos.

estados del nivel personal sino subpersonal. Se aplica a estados de subsistemas cognitivos que no entran en la explicación intencional de la conducta de los individuos puesto que ésta sólo apela a estados intencionales. Y, *a fortiori*, es irrelevante que haya algo subdoxástico compartido por los individuos a la hora de dar cuenta de sus comportamientos.

Por otro lado, la noción de contenido estrecho pensada para estados intencionales no es una noción que se sostenga por sí misma, requiere de la noción de contenido amplio. De esto me ocuparé a continuación pero antes quisiera adelantar mi conclusión. Si uno acepta lo desarrollado en §2, se sigue que la distinción estrecho-amplio no es fructífera en el nivel subpersonal porque en este nivel no hace falta distinguir aquel aspecto del contenido determinado por el mundo y aquel aspecto que tiene rol causal, básicamente porque no hay contenido amplio en los estados subpersonales.

Para mostrar esto es necesario concentrarse en enfoques efectivos de la noción de contenido estrecho y no en meras definiciones generales como he hecho hasta ahora. Desde que la noción surgió a la luz, es de utilización masiva pero muy pocos han desarrollado una teoría del contenido estrecho. Uno de ellos es Fodor (1987).<sup>41</sup> Por otro lado, hoy en día, los que siguen sosteniendo la idea del contenido estrecho son los que adhieren a la semántica del rol conceptual (Block 1986). Aunque a decir verdad nadie ofreció una explicación de la naturaleza del contenido estrecho,<sup>42</sup> dada la influencia de la semántica del rol conceptual en algunas áreas de la psicología cognitiva diré algo acerca de cómo se concibe la misma.<sup>43</sup>

---

<sup>41</sup> Fodor (1994) abandona la noción de contenido estrecho por considerarla superflua. Sin embargo, según Aydede (1997) la postura de Fodor anterior y posterior a (1994) es extensionalmente equivalente dado que desde un principio la noción de contenido estrecho colapsó con la de contenido amplio.

<sup>42</sup> Block (1995b, p. 174) dice explícitamente que no tiene una teoría acerca del contenido estrecho.

<sup>43</sup> Hay otros enfoques de la noción de contenido estrecho (véase Brown 2003 para una muy buena presentación de los diferentes enfoques). Pero casi todos convierten al contenido estrecho en derivativo del amplio (Cf. Brown 2003, Chalmers 2002, Stalnaker 1990). En los casos en los que aparentemente no es así, el contenido estrecho son las concepciones del agente (Chalmers 2002) o los mundos nocionales de los agentes (Dennett 1987, cap. 5), aspectos que, desde la perspectiva desarrollada en el capítulo anterior, pertenecen al nivel personal y por ende, no participan en los mecanismos y estados del nivel subpersonal. De manera que, si bien estos enfoques no caen bajo la segunda crítica (la dependencia del contenido estrecho del amplio), caen bajo la primera (la de ser enfoques para el nivel personal).

Según Fodor, lo que comparten los gemelos en la tierra y en la tierra gemela que tienen creencias acerca del agua es el contenido estrecho. El contenido estrecho es una función de contextos a condiciones de verdad (o contenido amplio). En el caso de un individuo que está en la tierra, Oscar<sub>1</sub>, la función proyecta este contexto en H<sub>2</sub>O, mientras que para el gemelo Oscar<sub>2</sub>, proyecta el contexto tierra gemela en XYZ. Los gemelos tienen idéntico contenido estrecho *agua* si comparten la misma función, esto es, si ciertos contrafácticos son verdaderos: si Oscar<sub>2</sub> estuviera en la tierra el concepto *agua* denotaría H<sub>2</sub>O y si Oscar<sub>1</sub> estuviera en la tierra gemela, *agua* denotaría XYZ. Así, el contenido estrecho de *agua* sería el conjunto de todas las relaciones mundos-propiedades, esto es el conjunto de todos los pares ordenados: {<tierra, H<sub>2</sub>O>, <tierra gemela, XYZ>, ...}.

Nótese que la identidad del contenido estrecho depende de ciertos contrafácticos, mencionados más arriba, que a su vez dependen de covariaciones nomológicas que conectan instancias de propiedades en el mundo con instancias de representaciones (ver capítulo 3, §1). El contenido amplio de los pensamientos acerca de agua de Oscar<sub>1</sub> está determinado por covariaciones nomológicas (puramente contrafácticas) entre instancias de H<sub>2</sub>O e instancias de *agua* (la idea sería: si se instanciara la propiedad H<sub>2</sub>O, causaría la instanciación del símbolo *agua*), en el caso de Oscar<sub>2</sub> está determinado por covariaciones nomológicas entre instancias de XYZ y de *agua* (o más específicamente de *agua gemela*). De manera que, en última instancia la identidad del contenido estrecho depende de las leyes que determinan (e individúan) el contenido amplio y si no hay tales leyes, entonces por bien definida que esté la noción de contenido estrecho, ésta sería inútil.

Así, los contenidos estrechos son constructos a partir de contenidos amplios. Fodor (1991, p. 300) lo dice de esta manera: “una condición necesaria para que un estado tenga contenido *estrecho* es que haya algún entorno en el que tuviera contenido amplio (algún contexto relativo al cual es semánticamente evaluable).” Estrictamente hablando, el contenido estrecho no es contenido, sino contenido potencial, para ser contenido tiene que estar anclado (determinado su valor semántico) y una vez anclado es contenido amplio. De manera que, en la versión de

---

Fodor, la noción de contenido estrecho requiere de una concepción de contenido amplio porque su individuación depende de covariaciones mente-mundo. Dado que para los estados subpersonales no hay ninguna especificación en términos de contenido amplio, la noción de contenido estrecho, sería inútil aplicada a los mismos.

Según la semántica del rol conceptual de doble factor el significado de una creencia (o concepto) depende, en parte, del rol que juega en la economía cognitiva del individuo, esto es, en la percepción, pensamiento, etc. y, en parte, de la referencia o condiciones de verdad.<sup>44</sup> Ambos aspectos son componentes del significado: el rol conceptual constituye el contenido estrecho (el aspecto no-relacional) mientras que la relación entre las representaciones y el mundo constituye el contenido amplio (el aspecto relacional). La semántica del rol conceptual sostiene que se especifica el contenido de una creencia una vez que se establece su rol conceptual y sus condiciones de verdad. El rol conceptual suele ser explicitado en términos funcionales y/o causales, así el rol funcional/causal de una creencia es el lugar que ocupa en la red inferencial/causal de las creencias. El contenido amplio toma la forma, en general, de una teoría causal de la referencia. La relación que hay entre los dos factores es que los roles conceptuales si bien determinan la naturaleza del factor referencial, no determinan la referencia en sí misma. En este sentido, los roles conceptuales que constituyen el contenido estrecho determinan una función que va de contextos a condiciones de verdad.

Sin embargo, no todas las propiedades inferenciales que constituyen un rol causal son pertinentes para la determinación de la función, aquellas que lo son sólo

---

<sup>44</sup> Se suele dividir el enfoque de la semántica de rol conceptual en teorías solipsistas y no-solipsistas y dentro de estas últimas se ubican las de doble factor y las de un solo factor. Cualquier semántica del rol conceptual solipsista (por ejemplo, una que sostenga que el contenido de las actitudes proposicionales se determina sólo por conexiones inferenciales) sería inviable porque no podría resolver los problemas por los cuales surgió la teoría causal de la referencia (por ejemplo, problemas de ignorancia y de representaciones erróneas) y estaría sujeta a argumentos conocidos acerca del lenguaje privado. Sería una concepción de contenido estrecho totalmente inviable.

Con respecto a las teorías de un solo factor como la de Harman (1982), éstas incluyen en los roles conceptuales 'de brazo largo' el aspecto referencial sin comprometerse con un contenido estrecho (Cf. Block 1986, p. 89) y en ese sentido no hay una concepción de contenido estrecho. Si la hubiera, dada la crítica de Block (1986), que sostiene que estas teorías pueden transformarse en teorías de doble factor, éstas estarían sujetas a la misma relativización del contenido estrecho con respecto al contenido amplio que las de doble factor.



constituyen una parte de los roles conceptuales que permiten dar cuenta de la conducta. El rol funcional de una creencia es importante a la hora de dar cuenta de la conducta porque ésta depende (en parte) de cómo los individuos representan los referentes. Así, si Edipo hubiera conocido la identidad Yocasta = mamá (quizá) no se hubiera casado con su madre. Evidentemente, *Yocasta* y *mamá* tenían roles inferenciales distintos en la economía cognitiva de Edipo. Cuando se habla de inferencias se lo hace en términos generales, éstas incluyen no sólo procesos deductivos sino asociaciones subjetivas, por ejemplo un individuo puede asociar ciertos conceptos de insectos con conceptos como *peligro*, mientras que otros no y estas asociaciones distintas dan lugar a comportamientos diferentes (aunque la asociación de cierto insecto con *peligro* no cumple un rol en la determinación de la referencia). Hay al menos cierto aspecto de la red inferencial que es accesible al sujeto y en este sentido los roles conceptuales son epistémicos. A diferencia de la noción fodoriana, que es puramente denotacional, la semántica del rol conceptual tiene un aspecto epistémico constitutivo del contenido estrecho.

Si bien el contenido estrecho constituye un factor (o un determinante) independiente del contenido intencional, depende del contenido amplio en el sentido de que es un término relativo a la noción de contenido amplio (como esta última noción es relativa a la de contenido estrecho dado que son dos determinantes del contenido, cosa que no ocurre en otros tipos de enfoques denotacionales, como el de Fodor, en los que la noción de contenido amplio se sostiene por sí misma). La noción de contenido estrecho es relativa a la de contenido amplio en el sentido de que los roles conceptuales determinan una función de contextos a condiciones de verdad (o contenido amplio). De manera que, los roles conceptuales no se sostienen por sí mismos para la determinación del contenido completo porque hay más para el contenido que lo que está en la cabeza de los individuos, de ahí que se requiera del contenido amplio. Según la semántica del rol conceptual, el contenido se descompone en dos factores, siendo uno de ellos el contenido estrecho y el otro, el factor referencial. Esta noción no es apropiada para los estados subpersonales dado que los contenidos de estos estados no requieren de un factor referencial.

A lo largo del capítulo, he hablado en términos de contenido suponiendo que los estados mentales en general poseen contenido de manera esencial y por ende la individuación de contenidos está basada en propiedades esenciales. En la sección siguiente abordaré este punto y además intentaré mostrar por qué la concepción del contenido interno de los estados subpersonales acuerda con los objetivos y el tipo de explicaciones en psicología cognitiva.

#### 4. Estados subpersonales y estados personales

Tanto en la teoría chosmkiana como en la de Marr, la concepción de los estados subpersonales está ligada a la de mecanismos computacionales. Siguiendo a Egan (1992, 1995) los estados computacionales no se individualizan en términos intencionales (esto es, en función del contenido especificado en términos del estímulo distal) sino formales. Las funciones computacionales son operaciones formales sobre estructuras simbólicas que se caracterizan en términos formales (por ejemplo, funciones matemáticas en el caso de la teoría de Marr). Las descripciones formales de un sistema computacional son independientes del entorno externo (dos mecanismos que computan la misma función son el mismo mecanismo aunque estén en dos entornos distintos) y del entorno interno (dos mecanismos idénticos podrían intercambiarse en un sistema sin alteración del mismo).

De manera que, “las descripciones computacionales son individualistas: individualizan los estados-tipo sin referencia al entorno del sujeto o su contexto social” (Egan 1992, p. 446). La implementación de funciones cognitivas en funciones computacionales es a través de una ‘función de interpretación’ que proyecta elementos simbólicos (representaciones) en elementos del dominio representado (Cummins 1989; Egan 1992, 1995). Es esta función de interpretación la que

permite asignar contenido semántico/intencional a los estados computacionales.<sup>45</sup> Dado que la propiedad de tener contenido intencional está dada por una función de interpretación, los estados computacionales no tienen contenido de manera esencial, o lo que es lo mismo el contenido de los estados computacionales no es esencial para la identidad de estos estados (como sí lo es su caracterización formal). Una misma función computacional podría tener diferentes interpretaciones y con ello diferentes contenidos.

Dado que los estados computacionales, que supervienen sobre estados físicos intrínsecos al sujeto, no se individualúan en función del contenido, y en ese sentido la caracterización intencional no le es esencial, es posible sostener un individualismo de estados mentales a la vez que un externalismo en cuanto a los contenidos. Esta es la postura de Egan (1995, 1999) quien sostiene que no hay ninguna motivación para adscribir contenido estrecho a los estados de los mecanismos computacionales subpersonales (como el visual). En cambio, habría consideraciones relacionadas con el rol explicativo del contenido que avalarían la adscripción de contenidos amplios. El argumento principal es que los *explananda* preteóricos de las teorías cognitivas se formulan en términos que hacen alusión al entorno del sujeto, de manera que el contenido tendría un rol explicativo en las teorías computacionales cognitivas que consiste en conectar las caracterizaciones formales de los sistemas con los *explananda* preteóricos. Así, según Egan, “es sólo bajo la interpretación de algunos de los estados del sistema como representaciones de propiedades distales ... que los procesos caracterizados ... por una teoría computacional se revelan como *visión*” (1995, p. 189).

Ahora bien, se podría discutir que la caracterización de los estados y procesos de una teoría psicológica se tenga que ajustar a la formulación preteórica de los *explananda* cognitivos en términos relacionales. Así como las categorías de las ciencias en general se apartan en su mayoría de las del sentido común, puede ocurrir que el estudio de los mecanismos cognitivos lleve a reformular los *explananda* de manera de que no se respeten sus caracterizaciones en términos

---

<sup>45</sup> Ver el capítulo 2 (§ 1.4). Mientras que Egan sostiene que la función de interpretación da lugar a caracterizaciones semánticas/intencionales, Cummins sostiene que da lugar a caracterizaciones sólo

que aluden al entorno del sujeto.<sup>46</sup> Si se acuerda con esto, no veo que haya una motivación explicativa de peso a favor del contenido amplio para sistemas cognitivos subpersonales. Pero, la cuestión que me parece importante discutir es la asimilación de la interpretación semántica a la intencional (o, lo que es lo mismo, del contenido interno al contenido intencional) porque es eso lo que hace que la discusión acerca de la individuación de los estados subpersonales se siga planteando en términos intencionales de contenido estrecho y amplio.

Hemos visto que los ESs no son estados intencionales de actitud proposicional (en el §1) y que no corresponde aplicarles las nociones de contenido intencional (estrecho y amplio) pensadas para las actitudes proposicionales (en §2 y §3). La razón fundamental para esto último es que la función de interpretación permite una interpretación semántica de los estados computacionales del sistema (esto es, una interpretación proposicional avalada por un análisis funcional, Cummins 1983) pero, para caracterizar a esos estados como intencionales se requiere además que las representaciones sean *para* el sistema (Cummins 1983, p. 78 y ss., Millikan 1989a, Loewer 1987). Si bien se podría pensar que puede haber una parte del sistema global para el cual los contenidos fueran significativos, creo que la opción más razonable es pensar que es el sistema global, la persona, para la cual las representaciones son significativas.

Las representaciones (el contenido interno) sobre las que operan los mecanismos subpersonales sólo llevan información que otros subsistemas pueden decodificar y lo hacen en función de las propiedades estructurales de las representaciones, no es información que esté disponible para la persona que es el sistema global. Son los estados personales (los intencionales de actitudes proposicionales y las experiencias perceptivas) a los que el sujeto tiene acceso, los que participan en el pensamiento y la acción, y en ese sentido los que significan

---

semánticas. Volveré sobre esto enseguida.

<sup>46</sup> Chomsky lo dice de esta manera: "Tampoco es relevante que la consideración de "representación" en entornos normales nos permita asociar el sistema bajo análisis con la función cognitiva de la visión, informalmente descripta. No es la tarea de la ciencia conformarse a las categorías de la intuición, o decidir si todavía es "visión" en entornos anormales....Por supuesto que el estudio de la percepción comienza con "tareas cognitivas" informalmente presentadas, pero se preocupa poco si algo similar a ellas se descubre a medida que el mismo progresa" (1995b, p. 56)

algo para el sujeto. Los estados intencionales no son acerca de los ESs sino acerca de los objetos o del mundo en general (Hamlyn, 1994; McDowell 1994b, Evans 1982, §7.4). Es en este sentido que "lo que el ojo de la rana le dice al cerebro de la rana no es lo que el ojo de la rana le dice a la rana" (Dennett 1977, p. 101).

El enfoque que surge de estas consideraciones refuerza la idea de que un sistema cognitivo tiene dos tipos de estados cognitivos. Los subpersonales para los cuales hay una caracterización semántica (contenido interno) y los estados personales para los cuales hay una caracterización intencional (contenido intencional). El contenido interno caracteriza los estados subpersonales en función de una interpretación semántica, avalada por un análisis funcional, que especifica en términos proposicionales la información que portan las estructuras de datos (representaciones).<sup>47</sup> Esta información sólo puede ser decodificada por otros mecanismos subpersonales, y en este sentido la persona no hace uso de ella para hablar acerca del mundo y realizar otras acciones. En cambio, el contenido intencional caracteriza los estados intencionales en función de una interpretación intencional que se especifica en términos que aluden a objetos en el entorno del individuo. La persona tiene acceso a dicho contenido, que participa en el pensamiento y la acción y por tanto, es contenido significativo para el individuo.<sup>48</sup>

Ahora bien, un objetivo razonable de la psicología y observable en la práctica psicológica es dar cuenta de cómo funcionan las capacidades cognitivas, esto es, dar

---

<sup>47</sup> La interpretación semántica está sujeta a un tipo similar de indeterminación que hemos visto en el capítulo anterior (§2) en relación a la interpretación intencional, pero por otros motivos, esto es, por la indeterminación que surge del enfoque computacional de la mente (Ver Cummins 1989, cap. 8; Hornstein 1991 y Egan 1995). Los estados de un sistema computacional no tienen una interpretación deliberada que da lugar a asignaciones unívocas de contenido porque hemos visto que el contenido no es esencial a los estados computacionales. Diferentes interpretaciones pueden ser compatibles con la función del sistema. Esto no quiere decir que el teórico pueda asignar cualquier interpretación, hay maneras de evitar interpretaciones no deseadas, pero siempre habrá lugar para la indeterminación en algunos casos. Por ejemplo, Cummins (1989, cap. 8) propone una manera de evitar las interpretaciones no deseadas que consiste en que la interpretación debe ser 'directa'. Es el sistema el que tiene que realizar la función y no la misma función de interpretación (o el intérprete). Este principio descarta, utilizando un ejemplo de Egan, "interpretar un escritorio como una calculadora dado que la asignación de números a los estados del escritorio requiere que el intérprete compute él mismo la función de adición" (1995, p. 193).

<sup>48</sup> Con respecto a la relación entre el nivel subpersonal y personal véase el capítulo 2 (§1.2 y nota 9). Por otro lado, una actitud más acorde a lo que se viene diciendo podría ser la de considerar al nivel subpersonal y personal estrictamente como dos niveles de explicación, dos relatos distintos acerca de los mecanismos y estados de los individuos. Volveré sobre esto en la conclusión.

cuenta de los mecanismos subpersonales y para ello, el contenido interno es todo lo que se requiere. La mayoría de los filósofos de la psicología creen que el objetivo de la psicología es dar cuenta de la conducta en función de explicaciones que apelan a contenidos intencionales. Por eso, en los primeros desarrollos de la filosofía de la psicología se analizaron exclusivamente los estados de actitudes proposicionales. Luego se vió que la psicología había alcanzado sus mayores logros dando cuenta del funcionamiento de capacidades cognitivas en las que intervienen estados cognitivos no-intencionales y se intentó dar cuenta de ellos en términos también intencionales. Esto es porque algunos creyeron que dado que son los estados intencionales a los que se apelan para dar cuenta de la conducta de los individuos, había que extender la noción de intencionalidad a aquellos estados subpersonales que participan en la cadena causal que lleva a la conducta.

Sin embargo, dados los objetivos razonables de la psicología cognitiva esta estrategia está mal encaminada, la psicología no tiene como objetivo las explicaciones intencionales de la conducta. En este sentido acuerdo plenamente con afirmaciones como las siguientes:

En la literatura filosófica, habitualmente se interpreta a la ciencia cognitiva como el estudio de cómo es causada la conducta por un complejo de creencias, deseos, etc. El enfoque del estudio de la mente recién esbozado [el estudio del lenguaje-I] no tiene nada que decir sobre esos temas. No estoy seguro de que sea un defecto dado que la ciencia cognitiva en ese sentido no existe realmente, y ni siquiera puede ser un objetivo razonable. No se conocen principios, ni siquiera imaginados, que van más allá de observaciones descriptivas de nivel inferior de limitada credibilidad y alcance. (Chomsky 1994b, p. 162). Si la "ciencia cognitiva" se entiende como una empresa interesada en la atribución intencional, podría resultar una actividad interesante (como lo es la literatura), pero probablemente no provea una teoría explicativa ni una que se integre en las ciencias naturales. (Chomsky 1992a, p. 23)

Gran parte de los filósofos de la psicología creen que una psicología científica sólo se puede lograr en la medida en que sean posibles las generalizaciones intencionales que apelan al contenido intencional de los estados

mentales. Sin embargo, las generalizaciones de la psicología cognitiva no pretenden dar cuenta de la complejidad de toda la acción humana, ni siquiera de la conducta aislada a la que subyacen ciertas capacidades. Ni la teoría lingüística ni la teoría de la visión intentan dar cuenta de la conducta lingüística ni la conducta visual, ni mucho menos de conductas aun más complejas en las que están involucradas las capacidades lingüísticas y visuales.

Las generalizaciones de la psicología pretenden dar cuenta de capacidades cognitivas y para ello apelan a estados internos y al funcionamiento de los mecanismos cognitivos. Son generalizaciones que tienen contenido empírico, se evalúan por investigación empírica, dan lugar a programas de investigación y están a la par de las generalizaciones de otras ramas de la ciencia. En cambio, las generalizaciones intencionales no parecen tener ninguna de estas propiedades (Cf. Chomsky 2000, pp. 165-66). Y la razón se debe, en gran parte a la imposición de un objetivo desmedido para la psicología, esto es, la explicación de la acción humana.

No estoy diciendo que la explicación del complejo comportamiento humano no sea un *desideratum* a tener en cuenta en la psicología cognitiva. Lo que creo es que la posibilidad de explicar esta complejidad no depende de cuán avanzada esté esta ciencia sino de la complejidad misma de la acción humana. La psicología cognitiva puede hacer su parte estudiando los sistemas subpersonales e incluso algunas de las interacciones entre los mismos, pero ella sola no será suficiente para cumplir el objetivo ambicioso. Seguramente se requerirán otras disciplinas para su explicación. Es en este sentido que si bien acuerdo con Chomsky en que la cuestión de la intencionalidad no forma parte de la psicología cognitiva, no acuerdo con sus razones, esto es, con la idea de que no puede ser objeto de estudio de las disciplinas empíricas y que, en última instancia, constituye un misterio.

Según Chomsky, un estudio naturalista “es una empresa humana particular que busca una clase especial de comprensión, obtenida por humanos en algunos pocos dominios cuando los problemas pueden simplificarse lo suficiente” (1995b, p. 10). Esa clase especial de comprensión es teórica y la otorgan las disciplinas que conforman las ciencias naturales. Son ellas las encargadas de dar cuenta de un ámbito bastante restringido de fenómenos pero cuya formulación es bastante precisa y cuyos

resultados pueden someterse a evaluación empírica. La lingüística en tanto parte de las ciencias naturales (es parte de la psicología y ésta en última instancia es parte de la biología) tiene como objetivo dar cuenta de los aspectos lingüísticos de los objetos naturales que somos los humanos, esto es, de los mecanismos y las representaciones mentales que constituyen la facultad del lenguaje.

Hemos visto que la facultad del lenguaje tiene dos componentes: un sistema cognitivo y sistemas de ejecución. La lingüística naturalista internalista se ocupa del sistema cognitivo, de su estado inicial (gramática universal) y su estado estable (lenguaje-I). Deja a un lado los sistemas de ejecución que constituyen el uso del lenguaje porque son de una complejidad extrema, en los que intervienen múltiples factores que no pueden ser aislados y simplificados para su estudio y en esa medida son poco proclives para una investigación naturalista (Chomsky 1991, p. 40). El uso del lenguaje está relacionado con las personas, con sus intereses, propósitos, sus puntos de vista, todas variables sumamente complejas (Chomsky 1992a, p. 22). En tanto que fenómenos como la intencionalidad están relacionados con los sistemas de ejecución, quedan fuera del alcance de una investigación empírica, esto es, una investigación naturalista internalista.

Esto concuerda perfectamente con la idea de que el ámbito de estudio más razonable para la psicología cognitiva es el subpersonal. Sin embargo, no creo que sea adecuado identificar a las ciencias empíricas con las ciencias naturales,<sup>49</sup> es posible que la cuestión de la intencionalidad pueda ser abordada por disciplinas empíricas como la antropología, la historia, y similares. En este sentido, las investigaciones empíricas de las ciencias naturales son internalistas mientras que las investigaciones empíricas naturalistas de otras disciplinas podrían aportar los aspectos no-internalistas necesarios para dar cuenta de la cuestión de la intencionalidad.

Chomsky (1995b, p. 28) no cree que esto sea posible. Las investigaciones naturalistas aunque no internalistas que toman en cuenta a los humanos desde el punto de vista de prácticas culturales, como participantes de comunidades, y demás no han dado frutos y quizá esto es así porque los fenómenos de los cuales intentan dar



cuenta estas investigaciones están más allá de nuestro alcance cognitivo, esto es, constituyen misterios.<sup>50</sup> Esto no significa que no podamos aprender acerca de los seres humanos como resultado de estas investigaciones, sólo que no aportan la comprensión teórica que otorgan las investigaciones naturalistas internalistas. Estas se ocupan de resolver problemas, esto es, cuestiones relacionadas con enfoques y conceptos bien comprendidos. De manera que, en una investigación empírica no queda otra opción que el internalismo (Chomsky 1995b, p. 49).

No acuerdo con la tesis mistericista acerca de los límites de nuestras capacidades cognitivas. Hasta donde sé, no hay un buen argumento que pueda establecer sobre bases empíricas que nuestra comprensión teórica tiene un límite, o, mucho menos, que pueda establecer la extensión precisa de nuestra comprensión teórica. El argumento del fracaso de las investigaciones naturalistas no-internalistas es altamente cuestionable. Se basa en la premisa de que todos los dominios en los que los problemas no pueden ser suficientemente simplificados están condenados al fracaso, esto es, no pueden aportar una comprensión teórica. Se podría acordar en que quizá se logra una comprensión teórica en el dominio de la ciencia natural porque los problemas pueden simplificarse por medio de una formulación precisa y enfoques y conceptos bien comprendidos, pero eso no significa que el resto de las ciencias empíricas no puedan alcanzar una comprensión teórica de otro tipo. Chomsky mismo sostiene que el naturalismo es más amplio que el internalismo, de modo que tener dos maneras de comprensión teórica no es una forma de dualismo metodológico.<sup>51</sup> En

---

<sup>49</sup> Si se asimila disciplinas empíricas a las ciencias naturales (como Chomsky parece hacer), dado que toda investigación de las ciencias naturales es internalista, entonces la intencionalidad no puede ser objeto de ninguna disciplina empírica.

<sup>50</sup> Para la diferencia entre problemas y misterios véase Chomsky (1975), cap. 4; (1994a), pp. 44-7. Según Chomsky, hay cuestiones relacionadas con la ejecución o el uso del lenguaje, como el 'aspecto creativo' que están más allá de nuestras capacidades cognitivas. Esto no es más que una parte de un problema general relacionado con la acción humana: el problema del libre albedrío en un mundo material. Es por eso que Chomsky (1991, pp. 40-1; 1992b, p. 72) siempre remarca que la conducta humana no está causada en ningún sentido que conozcamos del término. Las explicaciones intencionales (que intentan dar cuenta de la conducta apelando al rol causal de los estados intencionales) no sólo van más allá de una investigación internalista (y en ese sentido no forman parte de la psicología cognitiva) sino que son un caso perdido por el profundo aspecto mistericista que encierran.

<sup>51</sup> Chomsky (1995b) acusa de dualismo metodológico a quienes abocados a la consecución de un naturalismo metafísico tratan a la mente como si ésta no fuera un objeto de estudio natural.

este sentido, no hay razón para que haya sólo una manera de comprensión teórica naturalista.

Para concluir, en consonancia con el objetivo razonable de dar cuenta de cómo funcionan nuestras capacidades cognitivas, la explicación en psicología no consiste en explicaciones intencionales sino en la descomposición de una capacidad en subcapacidades más simples. Hemos visto en el capítulo anterior (§2) que la explicación psicológica suele adoptar la forma del análisis funcional de Cummins (1975, 1983). El análisis funcional encaja perfectamente en el tipo de explicación acerca del mecanismo del lenguaje-I. El mecanismo del lenguaje-I es un subsistema que participa del procesamiento del lenguaje (Chomsky 1975, p. 173; 1980a, pp. 214 y 237). Si de lo que se trata es de dar cuenta de la comprensión (capacidad analizada), la tarea del mecanismo computacional del lenguaje-I consiste en proyectar una descripción estructural de la emisión lingüística. Para ello, el sistema identifica las palabras, atribuye categorías léxicas y utiliza los principios parametrizados de la gramática universal para determinar la asociación entre las partes de la descripción (Chomsky, 1988, pp. 80 y 109).

Los estados cognitivos intermedios (de las subcapacidades) y el final (la descripción estructural) poseen contenido interno que se especifica por una interpretación semántica que consiste en una descripción proposicional de acuerdo al análisis funcional. Esta especificación es semántica porque se la realiza en términos proposicionales pero no es intencional porque en ella no se alude a estados individuados en función de algún aspecto del mundo externo a la persona. La información que se especifica por medio del análisis funcional es información a la que sólo acceden los otros subsistemas o componentes del sistema de esa capacidad. De manera que hay un tipo de explicación psicológica no-intencional, que es subpersonal pero no en el sentido de una explicación mecánica de estímulo-respuesta (Bérmudez 1995) ni puramente neurofisiológica (Dennett 1987, cap. 3) (ver capítulo 2, §1.1) sino una que hace uso de estructura mental y para la cual la noción de contenido interno es la más adecuada porque recoge el rol explicativo que poseen las representaciones para dar cuenta de los procesos internos del individuo.

## Conclusión

### Representaciones e intencionalidad: computación y comunicación

Voy a utilizar estas últimas páginas para ofrecer una visión de conjunto de la propuesta presentada, aclarar algunos puntos tratados en los capítulos anteriores, extender algunas cuestiones con algunas reflexiones quizás un tanto especulativas y marcar las diferencias de enfoque con algunas de las nociones y distinciones tradicionales que he utilizado en la tesis. A tal efecto, en primer lugar, volveré sobre el tema de la asimilación de la cuestión de la representación con la de la intencionalidad (§1). Diré algo más sobre la noción de representación que considero adecuada para el ámbito psicológico, haré algunas observaciones acerca del ámbito de la intencionalidad, y diferenciaré mi propuesta de otras ofrecidas en mi misma línea. En segundo lugar, me ocuparé de la distinción personal-subpersonal (§2). Aclararé en qué sentido la distinción que he utilizado se aparta de la tradicional y me ocuparé de algunas de las consecuencias que se derivan de mi versión de la distinción. En tercer lugar, haré algunos comentarios con respecto a posibles interpretaciones no deseadas de mi tesis principal acerca de la distinción entre el ámbito de la psicología cognitiva computacional y el ámbito de la comunicación (§3). Por último a partir de lo que se desprende de los capítulos anteriores, intentaré esbozar una respuesta al interrogante del título de la tesis (§4).

#### 1. Representación

### *1.1. Cuestiones filosóficas y psicológicas*

Gran parte de lo que considero que está equivocado en la corriente de la filosofía de la psicología que he criticado tiene sus raíces en la filosofía moderna, o al menos, en la interpretación que los filósofos han hecho de la filosofía moderna. La tradición filosófica suele considerar a los filósofos modernos como Descartes, Hume y, Locke como los padres del representacionalismo. Es indiscutible que hay motivos para considerar a Descartes como aquel que puso en primer plano a la mente e inauguró, por así decirlo, el estudio de la misma entendida como un espacio interior con sus contenidos. De ahí en más la mente ocupa un lugar central en la metafísica de la modernidad.

En particular, la noción de representación mental lleva en sí la carga de dar cuenta tanto de aspectos filosóficos como psicológicos. Desde Descartes hasta Kant, la cuestión psicológica del funcionamiento de la mente a través de las representaciones mentales es prioritaria tanto para dilucidar cuestiones metafísicas acerca de su naturaleza como para contestar preguntas epistemológicas acerca de cómo es posible nuestro conocimiento del mundo. Los filósofos contemporáneos agregaron, o desarrollaron en detalle, la cuestión semántica de cómo las representaciones se conectan con el mundo.

Así, apelar a las representaciones mentales permite, además de responder interrogantes psicológicos ligados a cómo funciona la mente, dar cuenta de cuestiones filosóficas relacionadas con aspectos metafísicos, epistemológicos y semánticos. Estos aspectos filosóficos están íntimamente relacionados en la medida en que lo que se supone que es el problema crucial que hay que explicar es cómo obtenemos conocimiento acerca del mundo y para ello se postulan representaciones de cierta naturaleza metafísica que están conectadas de cierta manera con el mundo.<sup>1</sup> Esta doble tarea psicológica y filosófica perdura en la

---

<sup>1</sup> Rorty (1979) desarrolla la cuestión de cómo la filosofía moderna asimiló el ámbito psicológico al epistemológico. Y muestra que no hay manera de que la psicología pueda tener ningún fin epistemológico, de modo de que no se pueden extraer consecuencias interesantes desde la explicación psicológica para la tarea epistemológica. Mi intención es mostrar, en paralelo con el planteo rortiano, que no hay manera de que la psicología pueda tener ningún fin semántico.

noción contemporánea de representación, tal como es entendida por la filosofía de la psicología, y es esto lo que considero que en gran parte motivó, por un lado, una concepción errónea de los objetivos, el alcance y el tipo de explicación de la psicología cognitiva, y por el otro, un abordaje inadecuado de la cuestión de la intencionalidad.

Si bien es cierto que en las teorías de los filósofos de la psicología dedicados a la semántica de las representaciones mentales, la cuestión epistemológica no interviene, como sí lo hace en otras teorías de filósofos de la mente interesados en el problema del escepticismo, cargar a la noción de representación con el peso de dar cuenta de la cuestión semántica de la conexión entre el individuo y el mundo lleva a una tensión en el interior de la psicología que no hace más que desvirtuar el objeto de estudio de la misma. Los filósofos embarcados en el PNC creen que las teorías de la psicología cognitiva requieren una noción de contenido intencional. Esto se debe, en parte, a que piensan que el objetivo de la psicología es dar cuenta del complejo comportamiento humano por medio de explicaciones intencionales que apelan a generalizaciones que relacionan contenidos intencionales. Puesto que de lo que se trata, para estos filósofos, es de dar cuenta de la acción descrita en términos intencionales, la explicación tiene que aludir al contenido de los estados intencionales como las creencias y deseos, contenido para el cual se requiere de una semántica que lo conecte con el mundo. De manera que, el asimilar la cuestión de la intencionalidad a la de la representación lleva a pensar que la psicología cognitiva requiere conectar sus estados representacionales con el mundo.

Uno de los objetivos de esta tesis es mostrar que hay que separar los aspectos filosóficos y psicológicos, y cuando se trata de hacer filosofía de la psicología lo razonable sería que las representaciones carguen sólo el peso psicológico, de lo contrario los problemas se incrementan sin que las soluciones aporten mucho a la comprensión de las cuestiones psicológicas. Teóricos cognitivos como Jackendoff (1992, cap. 8) también observan la tensión entre un aspecto filosófico y otro psicológico en la noción de representación de algunos filósofos de la psicología. Para este lingüista cognitivo no es tarea de la psicología preguntarse cuál es la relación entre la mente y el mundo, de manera que tenemos creencias

acerca de cosas en el mundo y éstas pueden ser verdaderas o falsas, sino cómo funciona la mente. En este sentido, la noción de representación no es intencional porque no es necesario que éstas sean acerca del mundo. Su ejemplo más utilizado es el de las representaciones fonológicas: ¿en qué sentido se puede decir que una representación fonológica es verdadera o falsa? Chomsky (1995b) también se pregunta irónicamente, tal como hemos visto en el capítulo anterior, por qué no hablar de un contenido fonético si se supone que las representaciones son acerca del mundo; si esto suena absurdo, ¿por qué no pensar también que lo es para el resto de las representaciones?

Tal como Fodor (1980) dijo cuando era un solipsista metodológico, nociones como verdad, referencia y las demás nociones semánticas no son categorías psicológicas, por eso no tienen ningún rol explicativo en las teorías psicológicas. Sin embargo, las propuestas para una psicología no-semántica/intencional no fueron satisfactorias. La exclusión de la noción de contenido intencional de las teorías psicológicas llevó a propuestas extremas como el eliminativismo de las nociones intencionales a favor de una psicología neurofisiológica (Rorty 1965, P. M. Churchland 1981, P. S. Churchland 1986) o a favor de una psicología sintáctica como la Teoría Sintáctica de la Mente (Stich 1983), o a propuestas menos extremas como la incorporación a la psicología de la noción de contenido estrecho (Block 1986, Fodor 1987).

La idea subyacente de estas propuestas es que las nociones intencionales de la PSC no encajan bien en la ciencia cognitiva. Mi propuesta de separar la cuestión filosófica de la intencionalidad de la cuestión psicológica de la representación va en esta misma línea pero se diferencia en un aspecto fundamental que hace de mi propuesta una alternativa viable a la psicología intencional mientras que las basadas en el eliminativismo o el contenido estrecho no lo son.<sup>2</sup> Dado que es en la propuesta eliminativista donde más se puede apreciar la diferencia con mi enfoque y dado que me he ocupado de la noción de contenido estrecho en el capítulo anterior, me concentraré en la primera. Los argumentos a favor del eliminativismo,

si bien son de distinta índole, en última instancia, aluden a la imposibilidad de que las categorías de la PSC sean proyectables y por ende, a la imposibilidad de que formen parte de la ciencia. Esto es, ante la imposibilidad de reducir explicativamente las categorías intencionales a las categorías científicas, se opta por desechar de la ciencia a las primeras.

El problema con esta estrategia es que básicamente asimila nociones semánticas a las intencionales y asimila ambas a nociones psicológicas. Así, se piensa que si nociones semánticas/intencionales/psicológicas como 'creencia' no participan de una teoría psicológica entonces la psicología debe ser neurofisiología o puro simbolismo formal (mera sintaxis). O, dicho de otra manera, si en la psicología no hay una noción de contenido descrita en términos intencionales entonces no hay contenido en absoluto, y si no hay contenido, no hay un nivel de descripción cognitivo/psicológico y, en última instancia, no hay psicología. Este es un falso dilema del cual me ocuparé en §2, aquí sólo me interesa remarcar que mi propuesta, en cambio, tiene un lugar para la noción de contenido y, en ese sentido, no propugna ni una psicología neurofisiológica ni una meramente sintáctica. El aspecto crucial es que este contenido, tal como se vió en el capítulo anterior y sobre el cual volveré en §1.2, no es intencional. En este sentido, si hay algo novedoso en mi propuesta en relación con la alternativa 'psicología intencional - no psicología en absoluto' es que, hasta donde sé, nadie propuso una separación completa del fenómeno de la intencionalidad del de la representación.

Dado el fracaso de las alternativas no-intencionales de psicología, la respuesta en vigencia de los filósofos de la psicología al falso dilema es la de optar por el otro cuerno, esto es, por una psicología intencional que asimila representación con intencionalidad. Así, el realismo intencional que carga a las representaciones mentales con propiedades intencionales lleva al desarrollo de una semántica de condiciones de verdad que le otorga a la noción de contenido intencional (amplio) un rol explicativo en psicología. Sin embargo, hemos visto que

---

<sup>2</sup> Para posiciones a favor de una psicología científica intencional ver Fodor (1987). Para críticas a la teoría sintáctica de la mente, ver, en particular, Aydede (en prensa) y Egan (1989). Para críticas a la noción de contenido estrecho, ver, entre otros, Stalnaker (1989, 1990).

esta no es la manera más adecuada de entender las representaciones que participan en las teorías psicológicas.

### *1.2. Representación y computación*

Las representaciones son aquello a lo que, desde un nivel abstracto de descripción, se les puede otorgar una interpretación semántica que consiste en una descripción proposicional de acuerdo a un análisis funcional computacional. Esta especificación es semántica porque se la realiza en términos proposicionales pero no es intencional porque en ella no se alude a ningún estado intencional especificado en términos de algún aspecto del mundo externo a la persona. La información que se especifica por medio del análisis funcional es información a la que sólo acceden los otros subsistemas o componentes del sistema de esa capacidad. Así, las representaciones son descripciones de elementos internos que portan información para que un sistema pueda funcionar, no tienen una semántica en el sentido de ser acerca de algo en el mundo sino que portan información necesaria para que los mecanismos funcionen, y en ese sentido, es información interna al sistema.

Hemos visto en el capítulo anterior que esta descripción del contenido interno de las representaciones es la más adecuada para dar cuenta del rol explicativo de las mismas en las teorías psicológicas. Dado que estas últimas se ocupan del funcionamiento de (sub)sistemas subpersonales, la noción de representación atinente no es la que proviene de la PSC pensada para el nivel personal intencional. Esto acuerda con ciertas reflexiones de psicólogos cognitivos, de distintas disciplinas, que consideran que:

Aunque las representaciones mentales son centrales a la psicología de sentido común pre-científica, la psicología de sentido común no provee una definición rigurosa de representación, mucho más de lo que la física de sentido común provee



una definición rigurosa de masa y energía. La [noción de] representación, definida rigurosamente, es un concepto matemático y computacional (Gallistel 2001, p. 2)<sup>3</sup>

Cuando utilizo palabras como “mente, “representación mental”, “computación mental” y otras similares, me refiero al nivel de caracterización abstracta de las propiedades de ciertos mecanismos físicos hasta ahora casi enteramente desconocidos ... De la misma manera se podría formular una teoría de la visión humana en términos concretos, refiriéndonos, por ejemplo, a las células específicas de la corteza visual y a sus propiedades; o se podría formular de manera abstracta en términos de ciertas formas de representación (por ejemplo imágenes o dibujos lineales), de computaciones sobre tales representaciones, de los principios organizativos que determinan la naturaleza de esas representaciones, etc. En este último caso, la indagación pertenecería propiamente al estudio de la mente, según la terminología que he adoptado.... (Chomsky 1980a, p. 13)

Tal como afirman las citas, la noción de representación es una noción técnica que pertenece claramente al ámbito computacional. No es una noción intencional tal como la que proviene de la PSC que establece una relación de correspondencia entre un elemento de la mente/cerebro y algo externo al sujeto. La noción computacional de representación alude a una caracterización abstracta, en el nivel del procesamiento computacional, de ciertos elementos de la mente/cerebro que portan información que es manipulada internamente por los mecanismos cognitivos. Dado que la información es interna al sistema, esto es, sólo ciertos subsistemas de la persona pueden decodificarla dado que está codificada en un código al que sólo acceden estos subsistemas subpersonales, la noción de contenido interno parece la más adecuada para dar cuenta del rol explicativo de las mismas en el funcionamiento cognitivo.

### *1.3. Representación y comunicación*

---

<sup>3</sup> La paginación es de la versión que aparece en la página web: [rucss.rutgers.edu/faculty/GnG/Mental\\_Reps.pdf](http://rucss.rutgers.edu/faculty/GnG/Mental_Reps.pdf)

Una vez que separamos la noción de representación del fenómeno de la intencionalidad, éste último se lo puede ubicar en el ámbito de las prácticas lingüísticas tal como algunos filósofos (por ejemplo, Rorty, Davidson) y científicos cognitivos (por ejemplo, Chomsky, Jackendoff) proponen. La intencionalidad, tal como hemos visto, parece depender de mecanismos que van más allá del enfoque de la psicología cognitiva. Hay un aspecto que tiene que ver con el punto de vista de los agentes, sus intereses, y demás, que interviene en la adscripción de creencias y similares, y este aspecto parece exceder una investigación centrada en los mecanismos cognitivos. Estas investigaciones se ocupan de (sub)sistemas cuyos outputs deben integrarse de manera de dar lugar al pensamiento y la acción. Es poco probable que esta integración pueda realizarse sólo haciendo uso de los recursos ontológicos, metodológicos, conceptuales y demás, de la psicología cognitiva.

Dar cuenta de la acción intencional de los humanos es una tarea sumamente compleja para la cual no hay actualmente ni principios generales ni generalizaciones plausibles ni, mucho menos, leyes *ceteris paribus*. Si la intencionalidad está en las prácticas comunicativas difícilmente pueda haber leyes acerca del comportamiento de la gente porque ésta puede agruparse sólo en casos excepcionales pero no cotidianamente: la gente hace cosas por muy diversos motivos. Se podrían hacer generalizaciones sobre aspectos muy rudimentarios de la conducta humana como el ejemplo habitual de que si las personas creen que hay fuego en un establecimiento y desean salir de él, *ceteris paribus*, harán lo necesario para salir. Pero, todo el comportamiento complejo de cada una de las personas quedaría fuera de estas generalizaciones. El problema, tal como hemos visto en el capítulo anterior, no se relaciona con el estado de la psicología como ciencia sino con el *explanandum* de la explicación intencional.

Dado que el *explanandum* es la acción intencional y dado que esto excede el estudio de la psicología cognitiva, algunos filósofos como van Gelder (1993, 1998) sostienen que la ciencia cognitiva no nos puede dar una completa comprensión de la mente y que una ciencia de la mente tiene que incluir las prácticas sociales. Mi propuesta sigue la línea de enfoques como este, en el sentido de que dar cuenta de los estados intencionales requiere un abordaje multidisciplinario relacionado con el

estudio de las prácticas comunicativas. Hemos visto en el capítulo 4 que lo que pensamos, creemos, y similares no es una cuestión de qué representaciones tenemos instanciadas en nuestros cerebros y con qué cosas del mundo físico se conectan, sino que la cuestión de la intencionalidad radica en que son las personas, que están inmersas en el mundo natural, las que a través de sus intereses, puntos de vistas y, en general de las prácticas comunicativas significan (/refieren) las expresiones lingüísticas que manifiestan creencias, deseos, y similares.

Nuestros estados intencionales lingüísticos se relacionan con el mundo por medio de relaciones complejas a explorar en las que participan, al menos, cuatro elementos: persona, expresión, mundo y contexto. Son las personas las que piensan algo por medio de una expresión en ciertas circunstancias. Las oraciones del lenguaje natural expresan los estados internos personales, pero esas oraciones no representan el mundo. Los estados internos personales como las creencias y deseos son estados que el sujeto puede manipular para la acción, son concientes y conceptualizados y en este último sentido, están expresados en oraciones del lenguaje natural. Pero estas oraciones no representan en el sentido técnico psicológico de llevar información a la que sólo los subsistemas cognitivos pueden acceder. La noción de representación es una noción técnica psicológica que se aplica a los estados internos subpersonales. Tampoco representan la realidad en el sentido de las representaciones intencionales del PNC. La intencionalidad no es para nada un fenómeno representacional, en el sentido de que no se trata de relacionar representaciones con el mundo porque, entre otras cosas:

...nadie descubrió una manera de individuar hechos o estados de cosas de un modo que ayude a explicar qué hecho representa una oración dada. Si al decir que el lenguaje representa hechos, sólo queremos decir que podemos usar oraciones para describir objetos y eventos, no se hace ningún daño. Esto es, después de todo, sólo una versión elaborada de la trivialidad de que algunas oraciones son verdaderas y algunas falsas. Pero, nos engañamos cuando decimos que el lenguaje representa la realidad (o cualquier otra cosa) a menos que podamos especificar de manera útil las entidades representadas. (Davidson 1997, p. 17)

El lenguaje no es un intermediario entre las personas y el mundo a la manera en que los defensores del PNC consideran a las representaciones intencionales. Así como no vemos el mundo a través de nuestras representaciones (psicológicas) no lo vemos a través de nuestro lenguaje. El lenguaje, siguiendo las reflexiones davidsonianas en este punto, es un órgano más con el que estamos directamente, aunque de una manera compleja, conectados con el mundo.

## 2. La distinción personal-subpersonal

En el §1 he ofrecido una visión de conjunto resumida de la propuesta de distinción entre los fenómenos de la representación y de la intencionalidad. Es innegable que esta idea tiene sus orígenes en otra idea que se puede expresar utilizando una frase ya clásica de M. Arbib: "lo que el ojo de la rana le dice al cerebro de la rana no es lo que el ojo de la rana le dice a la rana" (citada en Dennett 1977, p. 101). Lo que el ojo de la rana le dice al cerebro de la rana no es información a la que la rana accede, esto es, no es contenido significativo para ella. Esta información seguramente no está codificada en palabras pero se pueden construir oraciones que sean verdaderas acerca de lo que el ojo de la rana le dice a su cerebro (esta es toda la idea de la interpretación semántica del contenido interno). En cambio, lo que el ojo de la rana le dice a la rana es información a la que accede la rana, y yendo más allá de esta metáfora y cambiando su sujeto, lo que el ojo de la persona le dice a la persona es contenido significativo, muy probablemente conceptual, inferencial y conciente. La persona puede manipular esa información (no es algo que simplemente 'le ocurre') de manera que es la persona la que piensa y actúa en el mundo en función de esa información.

La psicología se ocupa de estados y mecanismos subdoxásticos subpersonales, no de los intencionales personales. De manera que el contenido de estos estados es interno a los subsistemas cognitivos. Los estados intencionales personales poseen contenido intencional que es el de las expresiones lingüísticas

que utilizamos en la comunicación para atribuir intencionalidad. Estos estados están conectados con el mundo porque el lenguaje conecta a las personas directamente con su entorno natural y social. El estudio de la atribución intencional en la comunicación lingüística es tarea que va más allá de la psicología y le corresponde a un conjunto de disciplinas empíricas. Se sigue de esto que no se requiere de una semántica intencional para los estados subpersonales ni para los personales. Por eso el PNC es superfluo.

Dicho de otra manera, mientras el PNC apunta a explicar la intencionalidad a partir de una teoría del contenido asimilando intencionalidad a representación, el PNA apunta a explicar la intencionalidad a partir de las prácticas comunicativas separando el fenómeno de la intencionalidad del de la representación. Mientras que el PNC, de alguna manera, asimila los niveles explicativos personal y subpersonal, el PNA los separa dejando para este último el ámbito representacional del cual se ocupa la psicología cognitiva. Así, la distinción personal-subpersonal juega un rol fundamental a lo largo de esta tesis. La distinción no es nueva como así tampoco la división de tareas que se funda en la misma. Dennett (1968, 1987) la introdujo con el fin de defender la idea de que la psicología cognitiva y su teoría de la actitud intencional pertenecen a niveles explicativos diferentes.

Sin embargo, si bien mi propuesta acuerda en general con la consecuencia de la distinción personal-subpersonal que se expresa en afirmaciones como las siguientes:

... las creencias y deseos no son los objetos adecuados de estudio de la psicología cognitiva. Dicho de otra manera, las teorías cognitivas son o deberían ser teorías del nivel subpersonal, en el que las creencias y deseos desaparecen, para ser reemplazados con representaciones de otros tipos sobre otros temas. (Dennett 1977, p. 105)

... la "teoría" del nivel personal acerca de las personas no es una teoría psicológica. (Dennett 1978 p. 154)

el contenido sustantivo de la distinción como así también de la consecuente división de tareas no concuerda con la versión tradicional de Dennett.

Dennett introdujo la distinción para enfatizar la idea de que ante la pregunta de cómo las personas distinguen la sensación de dolor de otras sensaciones, no hay nada que las personas hagan *qua* personas que permita responder a esta pregunta. Las personas simplemente distinguen las sensaciones de dolor, es un hecho bruto acerca de las mismas. Se pueden exigir explicaciones ulteriores de cómo las personas tienen la capacidad de discriminar dolor, pero entonces tendríamos que abandonar el nivel explicativo personal de los estados y procesos mentales para pasar al nivel subpersonal de los estados y procesos cerebrales. Una vez que pasamos al nivel subpersonal, la explicación es mecánica, en términos de movimientos del cuerpo humano u organización del sistema nervioso. Aquí ya se deja de hablar de dolor y otros estados mentales para hablar de procesos físicos. Si bien Dennett no descarta la pregunta por la relación entre ambos niveles, la deja abierta dado que considera que los fenómenos mentales (o el discurso de lo mental) no puede ponerse en correspondencia con los acontecimientos corporales (o el discurso corporal).<sup>4</sup>

La manera en que he utilizado la distinción se diferencia de lo que se acaba de exponer en, al menos, dos cuestiones. En primer lugar, la distinción alude a una demarcación entre el ámbito de la causalidad y el de la racionalidad o normatividad. El nivel personal es concebido, según el mismo Dennett, en la línea del análisis conceptual de Ryle y Wittgenstein con el aditamento de la actitud intencional. Hemos visto algo acerca de esta última en el capítulo 2. Sucintamente, la actitud intencional consiste en la adscripción de estados intencionales que un sistema cognitivo *debería* tener, dadas las circunstancias y los objetivos del mismo, y se predice que se comportará de manera racional, esto es, en virtud de tales estados intencionales. Las explicaciones personales se formulan en base a la actitud intencional, de manera que son explicaciones normativas. En cambio, según mi concepción de la distinción, el nivel personal de explicación de la acción intencional

---

<sup>4</sup> Dennett (1991b) es un intento por dar cuenta de la relación subpersonal-personal difícil de congeniar con la distinción categórica entre los niveles en Dennett (1968, 1987).

alude a la práctica lingüística de la adscripción intencional pero, tal como hemos visto en el capítulo 4, en las prácticas lingüísticas no hay ningún elemento normativo.

En segundo lugar, según la distinción tradicional, el nivel subpersonal no parece ser un nivel cognitivo, más bien parece ser un nivel neurofisiológico mecánico. Sin embargo, tal como hemos visto en el capítulo 5, según mi concepción del nivel subpersonal, éste es cognitivo. Hay descripciones en el nivel cognitivo o mental del funcionamiento de los subsistemas subpersonales en términos de representaciones y computaciones sobre las mismas. En este sentido, se puede hablar de procesos y mecanismos mentales en el nivel subpersonal. Lo que sí es cierto es que en este nivel subpersonal no intervienen los estados intencionales personales, de manera que las explicaciones no aluden a nociones como creencias y deseos. Retomando el falso dilema del que hablé en §1.1, no es cierto que si en la psicología no intervienen nociones intencionales, ésta no sería una ciencia distinta de la neurofisiología.

Todo lo que se requiere para que haya un nivel psicológico es un nivel intermedio de descripción de lo que hace un mecanismo entre el nivel de la PSC y el neurofisiológico. Tanto las descripciones neurofisiológicas como las psicológicas son subpersonales, la diferencia radica en que estas últimas están formuladas en vocabulario psicológico que alude a representaciones y mecanismos computacionales. Pero, el vocabulario psicológico no es el vocabulario intencional. En general, las teorías psicológicas sólo hacen uso de vocabulario intencional en presentaciones informales. Así, en varias ocasiones Chomsky, por ejemplo, utiliza términos como '*cognize*' o 'conocimiento del lenguaje' pero lo cierto es que no tienen ningún rol teórico en las explicaciones formales acerca de la facultad del lenguaje puesto que éstas son en términos de mecanismos computacionales y nociones técnicas como 'estado de la facultad del lenguaje', 'lenguaje-I', y demás (cf. Chomsky 2003).

Este nivel psicológico es el mismo al que se recurre cuando se intentan explicar los mecanismos de ciertos animales como el canto de algunas aves o el sistema de navegación de los insectos en términos de representaciones y mecanismos computacionales. De manera que el vocabulario psicológico se

relaciona más con la idea de la atribución de estructura interna a un sistema que permita dar cuenta de la interacción de ciertos componentes del mismo para lograr una función cognitiva que con la atribución de estados intencionales. El nivel de descripción psicológica puede a su vez ser un nivel mental si de lo que se trata es de dar cuenta de funciones mentales. En todo caso, así como no hay un límite preciso entre aquellos movimientos corporales que pueden tener una descripción intencional y en ese sentido, constituir una acción intencional (esto es, no hay un límite preciso para distinguir entre meras acciones físicas y acciones intencionales, cf. Dennett 1968, cap. 9),<sup>5</sup> no hay una línea claramente definida entre las funciones neurofisiológicas que pueden ser descritas desde un nivel más abstracto sólo en términos cognitivos (compartidos por aquellos sistemas que tienen mentalidad y los que no) y aquellas que pueden ser descritas en términos mentales (intencionales, cualitativos, emocionales y demás). El falso dilema puede plantearse, en parte, porque se asimilan una serie de nociones como 'intencional', 'cognitivo', 'mental', 'psicológico' y demás que, tal como se mencionó en varias ocasiones a lo largo de la tesis, convendría mantener distinguidas aunque no haya límites precisos.

Estas dos diferencias hacen que mi utilización de la distinción difiera de la formulación tradicional y en ese sentido, la división de tareas tiene un contenido sustancial distinto. La psicología se ocupa de un nivel subpersonal que es cognitivo/mental y el nivel personal es el de las prácticas lingüísticas que no son normativas. Con respecto a las relaciones entre ambos niveles admito que así como, hasta donde sé, nadie tiene una teoría sustantiva al respecto (ver nota 9 del capítulo 2 y nota 48 del capítulo 5), tampoco tengo mucho que decir al respecto. Creo que es una cuestión absolutamente empírica que se relaciona con el requisito de la integración cognitiva que hemos visto en el capítulo 2. Quizá resulte que no hay manera, tal como Dennett piensa, de que pueda ponerse en correspondencia biunívoca ambos niveles (ya sea desde una perspectiva ontológica o del discurso).

---

<sup>5</sup> Aunque véase el §2 del capítulo 4 acerca de la implausibilidad de que algo pueda ser descripto como 'meramente' físico.



Lo que sí me parecería adecuado sostener, desde mi interpretación de la distinción y más allá de cualquier postura sustantiva con respecto a la relación entre ambos niveles, es que tanto los estados subpersonales como los personales se desarrollan desde la perspectiva ontogenética, de manera conjunta. Los estados subpersonales se configuran por el sólo hecho de estar equipados biológicamente como estamos y los personales se configuran en función de que estamos inmersos al nacer en la prácticas lingüísticas. No parece adecuado sostener que primero vienen los estados subpersonales que dan lugar a los personales, de manera de que los primeros son los que compartimos con otras especies animales, sino que en los humanos, dadas nuestras capacidades conceptuales/lingüísticas, los estados personales se van configurando a la par de los subpersonales.

Voy a dejar aquí esta cuestión, dado que estas observaciones no son más que especulaciones en torno a un problema que, tal como se mencionó, es empírico, para abordar posibles interpretaciones no deseadas que pudieran surgir de mi tesis principal de la distinción entre el ámbito de la computación y la comunicación.

### 3. Computación y comunicación

El eje de la discusión a lo largo de los capítulos se centró absolutamente en una polémica en el ámbito de la filosofía de la psicología. En este sentido, he discutido las tesis básicas de lo que he denominado PNC. El PNC es un proyecto filosófico y mi ataque al mismo se centró en discusiones que tienen su origen en el ámbito filosófico. En contraposición al PNC, mi tesis central fue que la cuestión de las representaciones le corresponde a la psicología computacional mientras que la de la intencionalidad al ámbito de la comunicación. Esto podría dar lugar a una interpretación inadecuada de esta división de tareas. Sobre todo la idea de que la psicología cognitiva se ocupa de un nivel subpersonal de procesamiento de la información podría llevar a creer que mi intención fue delimitar de manera normativa el ámbito psicológico, esto es, establecer prescriptivamente de qué se

deben ocupar los psicólogos. Visto desde la perspectiva filosófica, también podría pensarse que mi intención fue prescribir la tarea de los filósofos de la psicología.

Sin embargo, nada está más alejado de mis objetivos que esta interpretación. Mi intención no fue establecer ninguna normativa acerca de lo que *debe* hacer o ser la psicología y la filosofía de la mente. Simplemente, intenté exponer y discutir un proyecto de la filosofía de la psicología que considero infructuoso tanto para dar cuenta de la noción de representación en el ámbito psicológico como para dar cuenta del fenómeno de la intencionalidad. La discusión del PNC se planteó en un plano puramente descriptivo. La idea fue describir en qué consiste este proyecto y mostrar por qué es infructuoso a partir de, por un lado, las dificultades que tiene para dar cuenta del fenómeno de la intencionalidad y por el otro, de la descripción de las teorías y la práctica de la psicología cognitiva computacional.

En este sentido, el énfasis estuvo puesto en la discusión de la concepción de la intencionalidad y de la psicología cognitiva del PNC, desde los problemas y el marco conceptual filosóficos. Dicho de otra manera, el objetivo de mi ataque fue una concepción filosófica que considera que la noción de representación intencional es clave para dar cuenta tanto del fenómeno de la intencionalidad como de las teorías psicológicas. Con respecto a esto último, la concepción de psicología de la que me he ocupado es la que la filosofía misma de la psicología dice ocuparse, esto es, la psicología cognitiva computacional. Es por ello que mi formulación de las tesis básicas de la psicología cognitiva, en el capítulo 1, se centró en este tipo de psicología. Así, se supone que la filosofía de la psicología está encargada, entre otras cosas, de abordar cuestiones conceptuales en relación con las nociones fundamentales de la psicología cognitiva computacional. Sin embargo, el PNC más que analizar la noción de representación que ofrecen las teorías psicológicas computacionales intenta imponer a las mismas la noción que considera que permite resolver el problema que al PNC le interesa resolver, esto es, el de la intencionalidad.

Como un productor de un disco que en vez de contribuir, por medio del análisis crítico, al mejoramiento de la propuesta musical del músico en cuestión, en cambio, intenta que el músico toque con los instrumentos, las partituras y los arreglos que el productor quiere, la filosofía de la psicología encarnada en el PNC intenta hacer

encajar las representaciones psicológicas en el molde que considera adecuado para resolver el problema de la intencionalidad. Así, intenta imponerle a la psicología cognitiva computacional un objetivo desmedido, esto es, la explicación de la acción intencional humana. Por otro lado, hemos visto, en particular en el capítulo 4, por qué la concepción de la intencionalidad del PNC no es fructífera. Mi intención no fue prescribir sobre estos temas sino mostrar por qué la concepción del PNC de la psicología computacional y de la intencionalidad es inadecuada en los términos mismos en los que el PNC planteó estas cuestiones.

#### 4. Representaciones mentales: ¿eslabón entre el individuo y el mundo?

A partir de lo que hemos visto se sigue que hay un sentido en que las representaciones son un eslabón entre el individuo y el mundo y un sentido en que no lo son. El sentido en el que no lo son creo que quedó claramente establecido. Según Fodor (1994, cap. 4) las representaciones mentales son intermediarios entre el mundo y los individuos, el mundo nos afecta indirectamente a través de las representaciones que son los constituyentes de los estados intencionales y actuamos en función de estas representaciones verídicas acerca del mundo. Es natural que en esta concepción surja la pregunta de cómo se relacionan las representaciones con el mundo.

Desde la perspectiva que he intentado defender, las representaciones en tanto categorías psicológicas no juegan ningún rol semántico en la conexión entre los estados mentales y el mundo sino que su rol es explicativo y radica en dar cuenta del funcionamiento de los mecanismos cognitivos internos al individuo. Estas representaciones del nivel subpersonal no requieren de una conexión con el mundo porque la conexión se da en el nivel personal de las atribuciones intencionales a través de las prácticas lingüísticas. Pero, el lenguaje no representa el mundo en ningún

sentido de ser un intermediario entre el individuo y su entorno sino que a la manera de un órgano más nos pone en contacto directo con el mundo.

Sin embargo, hay un sentido en que, al menos en los humanos y algunas otras especies, hay una estructura interna cognitiva representacional que media nuestra interacción con el entorno. Esta estructura interna es descripta desde el nivel de procesamiento de la información en términos de representaciones y computaciones. Pero hemos visto que estas representaciones no son intermediarias en el sentido de que *están en el lugar de algunas cosas del mundo*. Más bien son intermediarias en el sentido de que son *condiciones que hacen posible nuestra interacción con el entorno*. Así como estamos en contacto directo con otras personas cuando hablamos por teléfono a pesar de que utilizamos un intermediario que es el teléfono, estamos en contacto directo con nuestro entorno a través de nuestra estructura cognitiva. Ambos son medios que nos permiten interactuar, a pesar de que el primero, a diferencia del segundo, no es constitutivo de las personas. De manera que hay sentidos de 'contacto directo' que no son incompatibles con que haya 'mediaciones'.

El punto, entonces, no es si hay o no intermediarios entre las personas y su entorno. En general, tanto los filósofos de la psicología como los científicos cognitivos creen que hay cierto tipo de mediación,<sup>6</sup> la cuestión es cuál tipo de mediación es buena y cual no, cuál permite clarificar los interrogantes acerca de las capacidades cognitivas humanas y cuál no. Los defensores del PNC creen que postular objetos en la mente que son sustitutos de rasgos del mundo permite dar cuenta de las capacidades cognitivas y de la acción intencional. Mi propuesta del PNA sostiene que la postulación de objetos en la mente que son sustitutos de rasgos del mundo no ayuda sino que oscurece la explicación de estos fenómenos. Queda al arbitrio de la evidencia empírica y los argumentos conceptuales disponibles cuál tipo de intermediario es el mejor.

---

<sup>6</sup> Menos algunos enfoques en ciencia cognitiva como los de Brooks (1991) y la teoría de los sistemas dinámicos (Beer y Gallagher 1992, van Gelder 1995). Aunque véase Clark y Toribio (1994) para una defensa de la idea de que esas teorías están formuladas para dominios en los que, a decir verdad, no se requieren representaciones. No está claro que funcionen para dominios en los que sí se requerirían.

*Granny Laura:* bueno... ahora que terminamos esta exquisita delicia japonesa y antes de ver la película de la que les hablé, ¿me podrías aclarar la duda que tenía antes de cenar?

*Autora:* pero ¿cómo? ¿no escuchaste todo lo que acabo de decir?

*Granny Laura:* ¿dijiste algo importante?

*Auntie Carla:* dijo que las representaciones mentales no son un eslabón en el sentido en que Pedro lo cree sino en el que yo lo creo.

*Pedro:* bueno... por lo visto, todavía no está dicha la última palabra. Yo quiero tener mi derecho a réplica!

*Carla:* pero Pedro, vos también hablaste bastante durante la cena, ¿no te parece que ya dijiste lo tuyo?

*Greycat:* vamos todos! que la película ya empieza!

## Bibliografía

- Adams, F. y Aizawa, K. (1994) "Fodorian Semantics", en S. Stich y T. Warfield (eds.) *Mental Representation*, Oxford, Basil Blackwell, pp. 223-42.
- Anderson, J. (1980) *Cognitive Psychology and Its Implications*, San Francisco, Freeman.
- Armstrong, D. (1968) *A Materialist Theory of Mind*, London, Routledge & Kegan Paul.
- (1980) *The Nature of Mind*, Ithaca, NY, Cornell University Press.
- Aydede, M. (1997) "Has Fodor Really Changed His Mind on Narrow Content?", *Mind & Language* 12 (3/4): 442-58.
- (en prensa) "Computation and Functionalism: Syntactic Theory of Mind Revisited", en G. Irzik y G. Güzeldene (eds.) *Boston Studies in the History and Philosophy of Science*, Dordrecht, Kluwer.
- Barwise, J. y Perry, J. (1983) *Situations and Attitudes*, Cambridge, MA, MIT Press.
- Beer, R. y Gallagher, J. C. (1992) "Evolving Dynamical Neural Networks for Adaptive Behavior", *Adaptive Behavior* 1: 91-122.
- Bermúdez, J. L. (1995) "Nonconceptual Content: From Perceptual Experience to Subpersonal Computational States", *Mind & Language* 10 (4): 333-69.
- (2000) "Personal and Sub-personal: A Difference without a Distinction", *Philosophical Explorations* 3 (1): 63-82.
- Bilgrami, A. (1992) *Belief and Meaning*, Oxford, Blackwell.
- Block, N. (1978) "Troubles with Functionalism", en C. W. Savage (ed.) *Perception and Cognition. Issues in the Foundations of Psychology. Minnesota Studies in the Philosophy of Science* 9, Minneapolis, University of Minnesota Press, pp. 261-325. Reimpreso en N. Block (ed.) (1980) *Readings in the Philosophy of Psychology*, Vol. 1, Cambridge, MA, Harvard University Press, pp. 268-305.
- (1981) "What is Philosophy of Psychology?", en N. Block, *Readings in Philosophy of Psychology*, vol. 1, Cambridge, MA, Harvard University Press, pp. 1-8.
- (1986) "Advertisement for a Semantics for Psychology" en P. French, T. Uehling and H. Wettstein (eds.) *Midwest Studies in Philosophy*, Minneapolis, University of Minnesota Press, 10: 615-678. Reimpreso en S. Stich y T. Warfield (eds.) (1994) *Mental Representations*, Oxford, Basil Blackwell, pp. 81-141.
- (1995a) "On a Confusion about a Function of Consciousness", *Behavioral and Brain Sciences* 18: 227-87.
- (1995b) "Ruritania Revisited", *Philosophical Issues* 6: 171-87.

- Boghossian, P. (1989) "Review of Colin McGuinn's *Wittgenstein on Meaning*", *Philosophical Review* 1: 83-4.
- Botterill, G. y Carruthers, P. (1999) *The Philosophy of Psychology*, Cambridge, MA, Cambridge University Press.
- Brentano, F. (1874) *Psychologie vom Empirischen Standpunkt*, Leipzig. Trad. ingl.: *Psychology From An Empirical Standpoint*, New York, Humanities, 1973. [*Psicología*, Bs. As., Schapire, 1942]
- Broadbent, D. (1958) *Perception and Communication*, Oxford, Pergamon.
- Brooks, R. (1991) "Intelligence Without Representation", *Artificial Intelligence* 47: 139-59.
- Brown, C. (2003) "Narrow Mental Content", *The Stanford Encyclopedia of Philosophy*, URL= <http://plato.stanford.edu/archives/spr2003/entries/content-narrow/>.
- Bruner, J., Goodnow, J. y Austin, G. (1956) *A Study of Thinking*, New York, Wiley.
- Burge, T. (1986) "Individualism and Psychology", *Philosophical Review* 95: 3-46.
- Carnap, R. (1947) *Meaning and Necessity*, Chicago, Chicago University Press.
- Chalmers, D. (2002) "The Components of Content", en D. Chalmers (ed.) *Philosophy of Mind: Classical and Contemporary Readings*, Oxford, Oxford University Press, pp. 608-33.
- Chomsky, N. (1957) *Syntactic Structures*, The Hague, Mouton & Co.
- (1959) "Review of *Skinner's Verbal Behavior*", *Language*, 35: 26-58.
- (1965) *Aspects of the Theory of Syntax*, Cambridge, MA, MIT Press.
- (1975) *Reflections on Language*, New York, Pantheon. [*Reflexiones sobre el lenguaje*, Barcelona, Planeta Agostini, 1984.]
- (1980a) *Rules and Representations*, Oxford, Blackwell. [*Reglas y representaciones*, México, FCE, 1983.]
- (1980b) "Rules and Representations", *Behavioral and Brain Sciences* 3: 1-61.
- (1986) *Knowledge of Language*, New York, Praeger.
- (1988) *Language and Problems of Knowledge: The Managua Lectures*, Cambridge, MA, MIT Press. [*El lenguaje y los problemas del conocimiento*, Madrid, Visor, 1989.]
- (1991) "Linguistics and Cognitive Science: Problems and Mysteries", en A. Kasher (ed.) *The Chomskyan Turn*, Oxford, Basil Blackwell, pp. 26-53.
- (1992a) "Explaining Language Use", *Philosophical Topics* 20: 205-31. Reimpreso en N. Chomsky (2000) *New Horizons in the Study of Language and Mind*, Cambridge, MA, MIT Press, pp. 19-45.
- (1992b) "Language and Interpretation: Philosophical Reflections and Empirical Inquiry" en J. Earman (ed.) *Inference, Explanation, and Other Frustrations: Essays in the Philosophy of Science*, California, University of California, pp. 99-128. Reimpreso en N. Chomsky (2000) *New Horizons in the Study of Language and Mind*, Cambridge, MA, MIT Press, pp. 46-74.
- (1994a) "Naturalism and dualism in the study of language and mind" *International Journal of Philosophical Studies* 2: 181-200. Reimpreso en N.

- Chomsky *New Horizons in the Study of Language and Mind*, Cambridge, MA, MIT Press (2000a), pp. 75-105.
- (1994b) "Chomsky, Noam", en S. Guttenplan (ed.) *A Companion to The Philosophy of Mind*, Oxford, Blackwell, pp. 153-67.
- (1995a) *The Minimalist Program*, Cambridge, MA, MIT Press.
- (1995b) "Language and Nature", *Mind* 104 (413): 1-61.
- (1998) *Nuestro conocimiento del lenguaje humano*, Santiago de Chile, Universidad de Concepción & Bravo y Allende Editores.
- (2000) "Internalist Explorations" en N. Chomsky, *New Horizons in the Study of Language and Mind*, Cambridge, MA, MIT Press, pp. 164-94.
- (2003) "Replies", en L. Antony y N. Hornstein (eds.) *Chomsky and His Critics*, Oxford, Blackwell, pp. 255-328.
- Churchland P. M. (1981) "Eliminative Materialism and Propositional Attitudes", *Journal of Philosophy* 77: 67-90.
- Churchland, P. S. (1986) *Neurophilosophy*, Cambridge, MA, MIT Press.
- Clark, A. (1989) *Microcognition*, Cambridge, MA, MIT Press.
- Clark, A. y Toribio, J. (1994) "Doing Without Representing?", *Synthese* 101: 401-31.
- Collins, A. (1987) *The nature of Mental Things*, Notre Dame, Indiana, University of Notre Dame Press.
- Cummins, R. (1975) "Functional Analysis", *Journal of Philosophy* 72 (20) : 741-64.
- (1983) *The Nature of Psychological Explanation*, Cambridge, MA, MIT Press.
- (1986) "Inexplicit Information" en M. Brand y R. Harnish (eds.) *The Representation of Knowledge and Belief*, Tucson, The University of Arizona Press, pp. 116-26.
- (1989) *Meaning and Mental Representation*, Cambridge, MA, MIT Press.
- (1996) *Representation, Target and Attitude*, Cambridge, MA, MIT Press.
- (1997) "The Lot of the Causal Theory of Mental Content", *Journal of Philosophy* 94 (10): 535-42.
- Cussins, A. (1990) "The Connectionist Construction of Concepts", en M. Boden (ed.) *The Philosophy of Artificial Intelligence*, Oxford University Press, pp. 368-440.
- Davidson, D. (1968) "On Saying That", *Synthese* 19:130-46. Reimpreso en D. Davidson (1984) *Inquires into Truth and Interpretation*, Oxford, Clarendon Press. ["Al decir *that*", en *De la verdad y la interpretación*, Barcelona, Gedisa, 1995, pp. 108-22]
- (1970) "Mental Events", en L. Foster y J. Swanson (eds.) *Experience and Theory*, Amherst, University of Massachusetts Press, pp. 79-101. Reimpreso en D. Davidson (1980) *Essays on Actions and Events*, Oxford, Clarendon Press, pp. 207-27. ["Sucesos mentales" en *Ensayos sobre acciones y sucesos*, México, UNAM, 1995, pp. 263-87].
- (1984) "Communication and Convention", *Synthese* 59: 3-17. Reimpreso en (1984) *Inquiries into Truth and Interpretation*, Oxford, Oxford University Press, pp. 265-80 ["Comunicación y convención" en *De la verdad y de la interpretación*, Barcelona, Gedisa, 1990, pp. 263-76].



- (1986) "A Nice Derangement of Epitaphs", en R. Grandy and R. Warner (eds.) *Philosophical Grounds of Rationality*, Oxford, Oxford University Press, pp. 156-74. Reprinted en E. Lepore (ed.) (1986) *Truth and Interpretation. Perspectives on the Philosophy of Donald Davidson*, Oxford, Basil Blackwell, pp. 433-446.
- (1992) "The Second Person", en P. French and T. Uehling (eds.) *Midwest Studies in Philosophy 17: The Wittgenstein Legacy*. Notre Dame: Notre Dame University Press, pp. 255-267. Reimpreso en Davidson (2001) *Subjective, Intersubjective, Objective*, Oxford, Oxford University Press, pp. 107-21.
- (1997) "Seeing through Language", en J. Preston (ed.) *Thought and Language*, Cambridge, Cambridge University Press, pp. 15-27.
- (2001) "Comments on Karlovy Vary Papers", en P. Kotatko, P. Pagin y G. Segal (eds.) *Interpreting Davidson*, Stanford, CSLI, pp. 285-307.
- Davies, M. (1986) "Tacit Knowledge and the Structure of Thought and Language", en C. Travis (ed.) *Meaning and Interpretation*, Oxford, Blackwell, pp. 127-158.
- (1989) "Tacit Knowledge and Subdoxastic States", en A. George (ed.) *Reflections on Chomsky*, Oxford, Blackwell, pp. 131-52.
- (1991) "Individualism and Perceptual Content", *Mind*, 100 (4): 461-84.
- (1995) "Reply: Consciousness and the Varieties of Aboutness", en C. MacDonald y G. MacDonald (eds.) *Philosophy of Psychology. Debates on Psychological Explanation*, Oxford, Blackwell, pp. 356-92.
- (2000) "Persons and their Underpinning", *Philosophical Explorations* 3(1): 43-62.
- Dennett, D. (1969) *Content and Consciousness*, London, Routledge & Kegan Paul.
- (1977) "Critical Notice: *The Language of Thought* by Jerry Fodor", *Mind* 86(342): 265-280. Reimpreso con el título "A Cure for the Common Code?" en D. Dennett (1978), *Brainstorms*, Cambridge, MA, Bradford Books/MIT Press, pp. 90-108.
- (1978) *Brainstorms*, Cambridge, MA, Bradford Books/MIT Press.
- (1987) *The Intentional Stance*, Cambridge, MA, MIT Press. [*La actitud intencional*, Barcelona, Gedisa, 1991]
- (1991a) "Real Patterns", *Journal of Philosophy* 88 (1): 27-51. Reimpreso en Dennett (1998) *Brainchildren*, Cambridge, MA, MIT Press, pp. 95-120.
- (1991b) *Consciousness Explained*, New York, Little, Brown and Company. [*La conciencia explicada*, Barcelona, Paidós, 1995]
- (1993) "Back from the Drawing Board", en Dahlbom, B., *Dennett and His Critics*, Oxford, Basil, Blackwell, pp. 203-35.
- Dretske, F. (1980) "The Intentionality of Cognitive States", *Midwest Studies in Philosophy* 5: 281-94.
- (1981) *Knowledge and the Flow of Information*, Cambridge, MA, MIT Press. [*Conocimiento e información*, Barcelona, Salvat, 1987]
- (1986) "Misrepresentation" en R. Bogdan (ed.) *Belief, Form, Content, and Function*, Oxford, Clarendon Press, pp. 17-36.

- (1995) *Naturalizing the Mind*, Cambridge, MA, MIT Press.
- Dreyfus, H. (1979) *What Computers Still Can't Do*, Cambridge, MA, MIT Press.
- Egan, F. (1989) "What's Wrong with the Syntactic Theory of Mind", *Philosophy of Science* 56: 664-74.
- (1992) "Individualism, Computation and Perceptual Content", *Mind* 101: 443-59.
- (1995) "Computation and Content", *The Philosophical Review* 104 (2): 181-203.
- (1999) "In Defence of Narrow Mindedness", *Mind & Language* 14 (2): 177-94.
- Evans, G. (1981) "Semantic Theory and Tacit Knowledge" en S. Holtzmann y C. Leich (eds.) *Wittgenstein: To Follow a Rule*, Londres, Routledge and Kegan Paul, Londres. Reimpreso en G. Evans (1985) *Collected Papers*, Oxford, Oxford University Press, pp. 322-42.
- (1982) *The Varieties of Reference*, Oxford, Clarendon Press.
- Field, H. (1978) "Mental Representation", *Erkenntnis* 13 (1): 9-61. Reimpreso en S. Stich y T. Warfield (eds.) (1994) *Mental Representations*, Oxford, Basil Blackwell, pp. 34-77.
- Flanagan, O. (1984) *The Science of the Mind*, Cambridge, MA, MIT Press.
- Flanagan, O y Rey, G. (1998) "Skinner", en E. Craig (ed.) *Routledge Encyclopedia of Philosophy CD-Rom*, London.
- Fodor, J. (1968) *Psychological Explanation. An Introduction to the Philosophy of Psychology*, New York, Random House. [*La explicación psicológica. Introducción a la filosofía de la psicología*, Madrid, Cátedra]
- (1975) *The Language of Thought*, Nueva York, Harper & Row. [*El lenguaje del pensamiento*, Madrid, Alianza, 1984]
- (1978) "Propositional Attitudes", *The Monist* 61 (4). Reimpreso en J. Fodor (1981) *Representations*, Cambridge, MA, MIT Press, pp. 177-203. ["Las actitudes proposicionales", en E. Rabossi (comp.) *Filosofía de la Mente y Ciencia Cognitiva*, Barcelona, Paidós, 1995, pp. 173-203]
- (1980) "Methodological Solipsism Considered as Research Strategy in Cognitive Psychology", *The Behavioral and Brain Sciences* 3 (1): 63-109. Reimpreso en (1981) *Representations*, Cambridge, MA, MIT Press, pp. 225-253.
- (1983) *The Modularity of Mind*, Cambridge, MA, MIT Press. [*La modularidad de la mente*, Madrid, Morata, 1986]
- (1986) "Why Paramecia Don't Have mental Representations", *Midwest Studies in Philosophy* 10: 3-23.
- (1987) *Psychosemantics*, Cambridge, MA, MIT Press. [*Psicosemántica*, Madrid, Tecnos, 1994]
- (1990a) "A Theory of Content, I: The Problem" en J. Fodor *A Theory of Content and Other Essays*, Cambridge, MA, MIT Press, pp. 51-87.
- (1990b) "A Theory of Content, II: The Theory" en J. Fodor *A Theory of Content and Other Essays*, Cambridge, MA, MIT Press, pp. 89-136.
- (1990c) *A Theory of Content and Other Essays*, Cambridge, MA, MIT Press.

- (1990d) "Psychosemantics, Or Where Truth Conditions Came From", en W. Lycan (ed.) *Mind and Cognition*, vol. I, Cambridge, MA, Basil Blackwell, pp. 312-37.
- (1991) "Replies", en B. Loewer y G. Rey (eds.) *Fodor in Mind. Fodor and his Critics*, Oxford, Basil Blackwell, pp. 255-319.
- (1994) *The Elm and The Expert*, Cambridge, MA, MIT Press.
- (1998) *Concepts*, Oxford, Clarendon Press. [*Conceptos*, Barcelona, Gedisa, 1999]
- Fodor, J. y Lepore, E. (1992) *Holism. A Shopper's Guide*, Oxford, Basil Blackwell.
- Fodor, J. y McLaughlin, B. (1990) "Connectionism and the problem of systematicity: Why Smolensky's solution doesn't work", *Cognition* 35:183-204.
- Fodor, J. y Pylyshyn, Z. (1988) "Connectionism and Cognitive Architecture: A Critical Analysis", *Cognition* 28: 3-71.
- Gallistel, C. (1998) "Methodological and Scientific Behaviorism" en E. Craig (ed.) *Routledge Encyclopedia of Philosophy CD-Rom*, London.
- (2001) "Mental Representation, Psychology of", N. Smelser y P. Baltes (eds.) *International Encyclopedia of the Social and Behavioral Sciences*, New York, Elsevier.
- Garnham, A. (1994) "Cognitive Psychology" en S. Guttenplan (ed.) *A Companion to the Philosophy of Mind*, Oxford, Blackwell, pp. 167-76.
- Gomila Benejam, A. (1995) "Externalismo semántico y determinación del contenido: el enfoque teleológico de R. Millikan", *Análisis Filosófico* 15: 107-33.
- Guttenplan, S. (1994) "Holism", in S. Guttenplan (ed.) *A Companion to the Philosophy of Mind*, Oxford, Blackwell, pp. 347-8.
- Hacker, P. (1972) *Insight & Illusion*, Oxford University Press.
- Hamlyn, D. W. (1990) *In & Out of the Black Box*, Oxford, Basil Blackwell.
- (1994) "Perception, Sensation, and Non-Conceptual Content", *The Philosophical Quarterly* 44 (175): 139-53.
- Harman, G. (1982) "Conceptual Role Semantics", *Notre Dame Journal of Formal Logic* 23: 241-56.
- Haugeland, J. (1985) *Artificial Intelligence: The Very Idea*, Cambridge, MA, MIT Press. [*La Inteligencia Artificial*, México, Siglo XXI, 1988].
- (1991) "Representational Genera" en W. Ramsey, S. Stich y D. Rumelhart (eds.) *Philosophy and Connectionist Theory*, Hillsdale, NJ, Lawrence Erlbaum Associates, pp. 61-89.
- Hempel, C. y Oppenheim, P. (1948) "Studies in the Logic of Explanation", *Philosophy of Science*, 15: 135-175.
- Hinton, G., McClelland, J. Y Rumelhart, D. (1986) "Distributed Representations" en D. Rumelhart, J. McClelland and the PDP Research Group (eds.) *Parallel Distributed Processing*, Vol. 1, Cambridge, MA, MIT Press, pp. 77-109.
- Horgan, T. (1992) "From Cognitive Science to Folk Psychology: Computation, mental Representation, and Belief", *Philosophy and Phenomenological Research* 52 (2): 449-84.
- Hornstein, N. (1991) "Grammar, meaning, and Indeterminacy", en A. Hasher (ed.) *The Chomskyan Turn*, Oxford, Basil Blackwell, pp. 104-21.
- Hull, C. (1943) *Principles of Behavior*, New York, Appleton-Century-Crofts.

- Hull, C. et al. (1940) *Mathematico-Deductive Theory of Rote Learning*, New Haven, NJ, Yale University Press.
- Jackendoff, R. (1992) *Languages of the Mind*, Cambridge, MA, MIT Press.
- Jackson, F. (1982) "Epiphenomenal Qualia", *Philosophical Quarterly* 32:127-36.
- Kim, J. (1993) "Psychological Laws", en J. Kim *Supervenience and Mind*, Cambridge, Cambridge University Press, pp. 194-215.
- (1996) *Philosophy of Mind*, Boulder, Colorado, Westview Press.
- Kosslyn, S. (1980) *Image and Mind*, Cambridge, MA, Harvard University Press.
- Kosslyn, S. y Pomerantz, J. (1977) "Imagery, Propositions, and the Form of Internal Representations", *Cognitive Psychology* 9:52-76. Reimpreso en N. Block (ed.) (1981) *Readings in Philosophy of Psychology*, Vol 2, Cambridge, MA, Harvard University Press, pp. 150-69.
- Lewis, D. (1972) "Psychophysical and Theoretical Identifications", *Australasian Journal of Philosophy* 50 (3): 249-58.
- Loar, B. (1983) "Must Beliefs Be Sentences?", en P. Asquith y Nickles (eds.) *Proceedings of the Philosophy of Science Association for 1982*, Michigan, East Lansing.
- Loewer, B. (1987) "From Information to Intentionality", *Synthese* 70: 287-317.
- Lycan, W. (1981) "Form, Function and Feel", *Journal of Philosophy* 78: 24-49
- (1988) *Judgement and Justification*, Cambridge, Cambridge University Press.
- (1993) "A Deductive Argument for the Representational Theory of Thinking", *Mind and Language* 8 (3): 404-22.
- (1996) *Consciousness and Experience*, Cambridge, MA, MIT Press.
- (2003) "Chomsky on the Mind-Body Problem" en L. Antony y N. Hornstein (eds.) *Chomsky and His Critics*, Oxford, Blackwell, pp. 11-28.
- Manfredi, P. (1993) "Tacit Beliefs and Other Doxastic Attitudes", *Philosophia* 22 (1-2): 95-117.
- Marr, D. (1977) "Artificial Intelligence-A Personal View", *Artificial Intelligence* 9:37-48. Reimpreso en J. Haugeland (ed.) (1981) *Mind Design*, Cambridge, MA, MIT Press, pp. 129-142.
- (1982) *Vision*, San Francisco, Freeman. [*La visión*, Madrid, Alianza, 1985]
- Martin, M. (1994) "Perceptual Content", en S. Guttenplan (ed.) *A Companion To The Philosophy of Mind*, Oxford, Blackwell, pp. 463-71.
- McCulloch, W. y Pitts, W (1947) "A Logical Calculus of the Ideas Immanent in Nervous Activity", en W. McCulloch (1965) *Embodiments of Mind*, Cambridge, MA, MIT Press, pp. 19-39. Reimpreso en M. Boden (ed.) (1990) *The Philosophy of Artificial Intelligence*, Oxford, Oxford University Press, pp. 22-39.
- McDonald, C. y McDonald, G. (1995) *Connectionism. Debates on Psychological Explanation*, Oxford, Basil Blackwell.
- McDowell, J. (1994a) *Mind and World*, Cambridge, MA, Harvard University Press.
- (1994b) "The Content of Perceptual Experience", *The Philosophical Quarterly* 44 (175): 190-205.
- McGinn, C. (1989) *Mental Content*, Oxford, Basil Blackwell.
- Miller, G. (1956) "The Magical Number Seven, Plus o Minus Two: Some Limits on Our Capacity for Processing Information", *Psychological Review* 63: 81-97.

- Miller, G., Galanter, E. y Pribram, K. (1960) *Plans and the Structure of Behavior*, Nueva York, Holt, Rinehart & Winston.
- Millikan, R. (1984) *Language, Thought, And Other Biological Categories*, Cambridge, MA, MIT Press.
- (1986) "Thoughts Without Laws. Cognitive Science With Content", *The Philosophical Review* 95 (1): 47-80.
- (1989a) "Biosemantics", *The Journal of Philosophy* 86 (6): 281-297. Reimpreso en S. Stich y T. Warfield (eds.) (1994) *Mental Representations*, Oxford, Basil Blackwell, pp. 243-58.
- (1989b) "In Defense of Proper Functions", *Philosophy of Science* 56: 288-302.
- (1990) "Compare and Contrast Dretske, Fodor, and Millikan on Teleosemantics", *Philosophical Topics* 18 (2): 151-61.
- (1991) "Speaking up for Darwin", en B. Loewer y G. Rey (eds.) *Meaning in Mind. Fodor and his Critics*, Oxford, Basil Blackwell, pp. 151-64.
- (1993a) "On Mentalese Orthography", en B. Dahlbom (ed.) *Dennett and His Critics*, Oxford, Basil Blackwell, pp. 97-123.
- (1993b) "Explanation in Biopsychology", en J. Hail y A. Mele (eds.) *Mental Causation*, Oxford, Clarendon Press, pp. 211-32.
- (1995) "Reply: A Bet With Peacocke", en C. MacDonald y G. MacDonald (eds.) *Philosophy of Psychology: Debates on Psychological Explanation*, Oxford, Basil Blackwell, pp. 285-92.
- (1999a) "Wings, Spoons, Pills, And Quills: A Pluralist Theory Of Function", *The Journal of Philosophy* 96(4) : 191-206.
- (1999b) "Historical Kinds and the Special Sciences", *Philosophical Studies* 95 (1-2) : 45-65.
- (2000) "Naturalizing Intentionality", en *Philosophy of Mind. Proceedings of the Twentieth World Congress of Philosophy*, vol. 9 Charlottesville, Virginia, Philosophy Documentation Center.
- (2001) "What has Natural Information to do with Intentional Representation?", en D. Walsh (ed.) *Naturalism, Evolution and Mind*, Cambridge, Cambridge University Press, pp. 105-25.
- (2002) "Bio-functions: Two Paradigms", en R. Cummins, A. Ariew y M. Perlman (eds.) *Functions: New Readings in Philosophy of Psychology and Biology*, Oxford, Oxford University Press, pp. 113-43.
- Nagel, T. (1974) "What It Is Like to Be a Bat?", *The Philosophical Review* 83:435-50. Reimpreso en Nagel (1979) *Mortal Questions*, Cambridge, CUP, pp.165-80. ["Qué se siente ser murciélago" en *La muerte en cuestión*, México, FCE, pp. 256-77]
- Neisser, U. (1967) *Cognitive Psychology*, New York, Appleton-Century-Crofts.
- Newell, A. (1980) "Physical Symbol Systems", *Cognitive Science* 4: 135-83.
- Newell, A. y Simon H. (1972) *Human Problem Solving*, Englewood Cliffs, NJ, Prentice-Hall.
- Pavlov, I. (1927) *Conditioned Reflexes*, Oxford, Oxford University Press.
- Papineau, D. (1984) "Representation and Explanation", *Philosophy of Science* 51 (4): 550-72.

- (1987) *Reality and Representation*, Oxford, Blackwell.
- (1990) "Truth and Teleology", en D. Knowles (ed.) *Explanation*, Cambridge, Cambridge University Press, pp. 21-44.
- (1999) "Normativity and Judgement", *Proceedings of the Aristotelian Society*, *Supp.* 73: 17-43.
- Peacocke, C. (1986) *Thoughts: An Essay on Content*, Oxford, Basil Blackwell.
- (1992) *A Study of Concepts*, Cambridge, MA, MIT Press.
- (1994) "Content, Computation and Externalism", *Mind and Language* 9 (3):303-35. Reimpreso en E. Villanueva (ed.) (1995) *Philosophical Issues, Content*, 6: 227-64.
- (1995) "Concepts and Norms in a Natural World", en C. Macdonald y G. Macdonald (eds.) *Philosophy of Psychology. Debates on Psychological Explanation*, Oxford, Basil Blackwell, pp. 277-84.
- Pears, D. (1995) "Wittgenstein's Naturalism", *The Monist* 78 (4): 411-24.
- Perry, J. (1979) "The Problem of the Essential Indexical", *Nous* 13:3-21.
- Pettit, P. (1992) "The Nature of Naturalism", *Proceedings of the Aristotelian Society. Suppl.* 66 : 254-66.
- Pinker, S. (1994) *The Language Instinct*, Harmondsworth, Penguin.
- Putnam, H. (1960) "Minds and Machines", en S. Hook (ed.) *Dimensions of Mind*, New York, New York University Press. Reimpreso en A. R. Anderson (ed.) (1964) *Minds and Machines*, Englewood Cliffs, NJ, Prentice-Hall, pp. 72-97. ["Mentes y máquinas", en A. R. Anderson (ed.) *Controversia sobre mentes y máquinas*, Barcelona, Tusquets, 1984, pp. 113-49]
- (1967) "Psychological Predicates" en W. Capitan y D. Merrill (eds.) *Art, Mind, and Religion*, Pittsburgh, PA, University of Pittsburgh Press, pp. 37-48. Reimpreso en Putnam (1975) *Mind, Language and Reality: Philosophical Papers*, Vol. 2, Cambridge, Cambridge University Press, pp. 429-40. ["La naturaleza de los estados mentales", *Cuadernos de Crítica*, UNAM, 1981]
- (1975) "The Meannig of "Meaning"", en K. Gunderson (ed.) *Language, Mind and Knowlegde*, Minneapolis, University of Minnesota Press, pp. 131-93.
- (1992) *Renewing Philosophy*, Cambridge, MA, Harvard University Press. [*Cómo renovar la filosofía*, Madrid, Cátedra, 1994]
- Pylyshyn, Z. (1980) "Computation and Cognition: Issues in the Foundations of Cognitive Science", *The Behavioral and Brain Sciences* 3:111-69.
- (1981a) "Imagery an Artificial Intelligence" en N. Block (ed.) *Readings in Philosophy of Psychology*, Vol 2, Cambridge, MA, Harvard University Press, pp:170-94.
- (1981b) "The Imagery Debate: Analog Media versus Tacit Knowledge", *Psychological Review* 88:16-45.
- (1984) *Computation and Cognition*, Cambridge, MA, MIT Press.
- Quesada, D. (1995) "La concepción teleológica de los estados mentales y de su contenido", en F. Broncano *La mente humana*, Madrid, Trotta, pp. 77-95.
- Quine (1956) "Quantifiers and Propositional Attitudes", *Journal of Philosophy* 53. Reimpreso en (1966) *The Ways of Paradox and Other Essays*, New York, Random House. ["Cuantificadores y actitudes proposicionales" en T. Simpson

- (ed.) (1973) *Semántica filosófica: problemas y discusiones*, Buenos Aires, Siglo XXI, pp. 217-33]
- (1960) *World and Object*, Cambridge, MA, MIT Press.
- (1975) "Five Milestones of Empiricism" en Quine, W. V. (1981) *Theories and Things*, Cambridge, MA, Harvard University Press, pp. 67-72. ["Cinco hitos del empirismo", en *Teorías y cosas*, México, UNAM, 1986, pp. 87-93]
- (1990) "Three Indeterminacies", en R. Barret y R. Gibson (eds.) *Perspectives on Quine*, Oxford, Basil Blackwell, pp. 1-16.
- Ramsey, W.; Stich, S. y Garon, J. (1990) "Connectionism, Eliminativism and the Future of Folk Psychology" *Philosophical Perspectives* 4:499-533. Reimpreso en W. Ramsey, S. Stich y D. Rumelhart (eds.) (1991) *Philosophy and Connectionist Theory*, Hillsdale, NJ, LEA, pp. 199-228.
- Reber, A. (1987) "The Rise and (Surprisingly Rapid) Fall of Psycholinguistics", *Synthese* 72: 325-339.
- Rey, G. (1991) "An Explanatory Budget for Connectionism and Eliminativism" en T. Horgan y J. Tienson (eds.) *Connectionism and the Philosophy of Mind*, Dordrecht, Kluwer Academic Publishers, pp. 219-240.
- (2003) "Chomsky, Intentionality, and CRTT", en L. Antony y N. Hornstein (eds.) *Chomsky and His Critics*, Oxford, Blackwell, pp. 105-39.
- Richard, M. (1994) "What Isn't a Belief?", *Philosophical Topics*, 22 (1 & 2): 291-318.
- Rorty, R. (1965) "Mind-Body Identity, Privacy and Categories", *Review of Metaphysics* 19(1): 45-8.
- (1979) *Philosophy and the Mirror of Nature*, Princeton, NJ, Princeton University Press. [*La filosofía y el espejo de la naturaleza*, Madrid, Cátedra, 1989]
- Rumelhart, J. y McClelland, J. and the PDP Research Group (1986) *Parallel Distributed Processing*, Vol. 1 y 2, Cambridge, MA, MIT Press.
- Russel, B. (1940) *An Inquiry into Meaning and Truth*, London, George Allen & Unwin. [*Significado y verdad*, Barcelona, Ariel, 1983]
- Searle, J. (1980) "Minds, Brains and Programs", *Behavioral and Brain Sciences* 3:417-24. Reimpreso en D. Hofstadter y D. Dennett (eds.) (1981) *The Mind's I: Fantasies and Reflections on Self and Soul*, New York, Bantam Books, pp. 353-72. [*El ojo de la mente*, Buenos Aires, Sudamericana, 1983, pp. 454-493]
- (1990) "Consciousness, Explanatory Invention and Cognitive Science". Reimpreso en C. MacDonald y G. MacDonald (eds.) (1995) *Philosophy of Psychology*, Oxford, Basil Blackwell, pp. 331-55.
- Segal, G. (1989) "Seeing What Is Not There", *Philosophical Review* 98: 189-214.
- (1991) "Defence of a Reasonable Individualism", *Mind* 100 (4):485-94.
- Schiffer, S. (1981) "Truth and the Theory of Content" en H. Parret y J. Bouvaresse (eds.) *Meaning and Understanding*, Berlin, Walter de Gruyter, pp. 204-22.
- (1991) "Ceteris Paribus Laws", *Mind* 100: 1-17.
- Shannon, C. (1948) "The Mathematical Theory of Communication", *Bell System Technical Journal* 27 (July and October): 379-423 y 623-656. Reimpreso como Shannon, C. y Weaver, W. (1949) *The Mathematical Theory of*

- Communication*, Urbana, IL, University of Illinois Press. [*Teoría matemática de la comunicación*, Forja, Madrid, 1981]
- Skidelsky, L. (2002) "Teorías del contenido mental o una naturalización cientificista", en D. Pérez (comp.) *Los caminos del naturalismo*, Buenos Aires, Eudeba, pp. 131-69.
- (en prensa) "Mental Content: Many Semantics, One Single Project", *Diálogos* 82.
- Skinner, B. (1938) *The Behavior of Organism: An Experimental Analysis*, New York, Appleton-Century-Crofts.
- (1953) *Science and Human Behaviour*, New York, MacMillan & Co.
- (1957) *Verbal Behavior*, New York, Appleton-Century-Crofts.
- Smart, J. (1959) "Sensations and Brain Processes", *Philosophical Review* 68:141-56.
- Smith, N. (1999) *Chomsky: Ideas and Ideals*, Cambridge, Cambridge University Press.
- Smolensky, P. (1988) "On the Proper Treatment of Connectionism", *The Behavioral and Brain Science*, 11:1-74.
- (1991) "Connectionism, Constituency, and the Language of Thought", en B. Lower y G. Rey (eds.) *Meaning in Mind. Fodor and His Critics*, Oxford, Basil Blackwell, pp. 201-227.
- (1994) "Computational Models of Mind", en S. Guttenplan (ed.) *A Companion to the Philosophy of Mind*, Oxford, Blackwell, pp. 176-85.
- Stalnaker, R. (1976) "Propositions" en A. MacKey y D. Merrill (eds.) *Issues in the Philosophy of Language*, Londres, Routledge and Kegan Paul, pp. 79-91.
- (1984) *Inquiry*, Cambridge, MA, MIT Press.
- (1989) "On What's in the Head", *Philosophical Perspectives* 3: 287-316.
- (1990) "Narrow Content", en C. Anderson y J. Owens (eds.) *Propositional Attitudes: The Role of Content in Logic, Language, and Mind*, Stanford, CSLI, pp. 131-45. Reimpreso en R. Stalnaker (1999) *Context and Content*, Oxford, Oxford University Press, pp. 194-209.
- Sterelny, K. (1990) *The Representational Theory of Mind*, Oxford, Basil Blackwell.
- Stich, S. (1978) "Beliefs and Subdoxastic States", *Philosophy of Science* 45: 499-518.
- (1983) *From Folk Psychology to Cognitive Science*, Cambridge, MA, MIT Press.
- (1996) *Deconstructing the Mind*, Oxford, Oxford University Press.
- Stone, T. y Davies, M. (2002) "Chomsky Amongst Philosophers", *Mind & Language* 17 (3): 276-89.
- Thagard, P. (2003) "Cognitive Science", *The Stanford Encyclopedia of Philosophy* (Spring Edition), URL= <<http://plato.stanford.edu/archives/spr2003/entries/cognitive-science/>>.
- Turing, A. (1936) "On Computable Numbers, With An Application to The Entscheidungsproblem", *Proceedings of the London Mathematical Society* 2 (42): 230-65; correction *ibid* 43: 544-46, 1937. Reimpreso en M. Davies (ed.) (1965) *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvability Problems and Computable Functions*, New York, Raven Press, pp. 115-54.



- (1950) "Computing Machinery And Intelligence", *Mind* 59 (2236): 433-60. Reimpreso en A. R. Anderson (ed.) (1964) *Minds and Machines*, Englewood Cliffs, NJ, Prentice-Hall, pp. 4-30. ["Maquinaria computadora e inteligencia", en A. R. Anderson (ed.) *Controversia sobre mentes y máquinas*, Barcelona, Tusquets, 1984, pp. 11-50]
- Tye, M. (1992) "Naturalism and the Mental", *Mind* 101 (403) : 421-41.
- (1996) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*, Cambridge, MA, MIT Press.
- van Gelder, T. (1991) "What is the "D" in "PDP"? A Survey of the Concep of Distribution" en W. Ramsey, S. Stich y D. Rumelhart (eds.) *Philosophy and Connectionist Theory*, Hillsdale, NJ, LEA, pp. 33-59.
- (1993) "The Distinction Between Mind and Cognition", en Y-H. Houng (ed.) *Mind and Copgnition: Collected Papers from the 1993 International Symposium on Mind and Cognition*, Taipei, Academia Sinica, pp. 1-29.
- (1995) "What Might Coguition Be If Not Computation?", *Journal of Philosophy*, 91 (7): 345-381.
- (1998) "Monism, Dualism, Pluralism", *Mind & Language* 13 (1): 76-97.
- Warfield, T. y Stich, S. (1994) "Introduction", en S. Stich y T. Warfield (eds.) *Mental Representations*, Oxford, Basil Blackwell, pp. 3-8.
- Watson, J. (1913) "Psychology As The Behaviorist Views It", *Psychological Review* 20: 158-77.
- (1925) *Behaviorism*, New York, Mouton.
- Wittgenstein, L (1921) *Tractatus Logico-Philosophicus* publicado como "Logisch-philosophische Abhandlung", *Annalen der Naturphilosophie*, 14: 185-262. Trad. inglés por C. K. Ogden *Tractatus Logico-Philosophicus*, Londres, Routledge & Kegan Paul, 1922. Nueva trad. inglés por D. F. Pears and B. F. McGuinness, London, Routledge and Kegan Paul, 1961. [*Tractatus Logico-Philosophicus*, Madrid, Alianza, 1973]
- (1953) *Philosophische Untersuchungen / Philosophical Investigations*, trad. G. E. M. Anscombe, Oxford, Blackwell. Reedición 3ra. (1958) *Philosophical Investigation*, Oxford, Basil Blackwell.
- Woodfield, A. (1986) "Two Categories of Content", *Mind & Language* 1 (4): 319-54.
- (1990) "The Emergence of Natural Representations", *Philosophical Topics* 18 (2): 187-213.

DIRECCIÓN DE BIBLIOTECAS  
 CENTRO DE INVESTIGACIONES Y ESTUDIOS  
 UNIVERSIDAD DE BUENOS AIRES