

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE FILOSOFÍA Y LETRAS
INSTITUTO DE INVESTIGACIONES FILOSÓFICAS

"REPRESENTACIÓN MENTAL, NATURALISMO Y
TELEONOMÍA:
ANÁLISIS DE UNA PROPUESTA TELEOSEMÁNTICA"

TESIS QUE PARA OPTAR POR EL GRADO DE MAESTRO EN FILOSOFÍA DE LA
CIENCIA PRESENTA

FERMÍN FULDA DE LA GARZA

DIRECTORA: MAITE EZCURDIA OLIVARRIETA

OCTUBRE, 2007.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos.

Esta tesis esta dedicada a mis padres, Irma de la Garza y Fermín Fulda por su el apoyo y estímulo que me han brindado en este andar por el camino de la filosofía.

Agradezco a todos los maestros que me instruyeron en la filosofía de la ciencia y en la filosofía en general. En particular la Dra. Maite Ezcurdia, la Dra. Ana Rosa Pérez, Dra. Ángeles Eraña, Dra. Salma Saab, el Dr. Raymundo Morado, el Dr. Sergio Martínez, Dr. Carlos López, Dr. Silvio Pinto y muy especialmente el Dr. Rasmus Winther, maestro y amigo.

Dedico esta tesis también a mis compañeros de la maestría, por haberme abierto a distintos intereses y perspectivas. En particular a Carla Merino, Fernanda Samaniego, José Ricardo Sánchez, Rodrigo Munguia, Alejandro Rangel y Fabricio McManus. Y a mis amigos, Oscar Espinosa, Andrés Figueroa, Iñigo Ahedo, Jonathan Camargo, Luis Rodrigo Basaguren, Manolo Fernández y especialmente a mi hermano Andrés Fulda, por las constantes y estimulantes conversaciones filosóficas.

Agradezco también al CONACYT y al DGEP por su apoyo con la beca que recibí durante el periodo de 2004 al 2006.

INDICE

INTRODUCCIÓN: Naturalismo Evolucionista.....	1
1. “Humano, demasiado humano”.....	1
2. Naturalismo y Mente.....	5
3. Contenido y representación mental.....	11
4. Teleosemántica.....	18
CAPÍTULO PRIMERO: Teleología Natural, Funciones Biológicas y Teoría Etiológica.....	22
1. Introducción: teleología natural.....	22
2. Explicación teleológica y teoría (proto)etiológica.....	25
3. Teoría etiológica histórica.....	27
4. Teoría Propensionista.....	32
5. Teoría de la historia moderna.....	46
6. Conclusión: Funciones de rol causal y el pluralismo des-unificador.....	49
CAPÍTULO SEGUNDO: Representación Mental y Teleonomía.....	52
1. Introducción:.....	52
2. El argumento teleosemántico.....	58
3. ¿Qué son y cómo representan las representaciones mentales?.....	64
a) Las funciones propias de la cognición.....	64
b) Mapas o íconos intencionales.....	66
c) Creencias y deseos.....	70
4. ¿Cómo pueden ser falsas las representaciones mentales?.....	73
5. Conclusión: consecuencias metafísicas del naturalismo Darwinista.	77

CAPÍTULO TRES: La Objeción de la Indeterminación y su Disolución	83
1. Introducción:.....	83
2. La objeción.....	84
3. La opacidad causal de los contextos selectivos.....	88
4. La jerarquía niveles de atribución funcional y la falacia del argumento del objeto de selección.....	92
5. Niveles teóricos y explicación intencional: el debate por el beneficiario	97
6. Conclusión: grados de determinación, “actos de identificación” y pluralismo..	103
CONCLUSIONES.....	107
APÉNDICE.....	111
BIBLIOGRAFÍA.....	112

INTRODUCCIÓN:

Naturalismo Evolucionista

“No account of the universe in terms merely of the redistribution of matter in motion is complete, no matter how true as far as it goes, for it ignores the cardinal fact that the character of matter in motion and of its redistribution is such as cumulatively to achieve ends –to effect the world of values we know. Deny this and you deny evolution; admit it and you admit purpose in the only objective –that is, the only intelligible- sense of that term. I do not say that in addition to the mechanism there are other ideal causes or factors which intervene. I only insist that the whole story be told, that the character of the mechanism be noted –namely, that it is such as to produce and sustain good in a multiplicity of forms. Mechanism is the mechanism of achieving results. To ignore this is to refuse to open our eyes to the total aspects of existence.”

John Dewey, 1910.

1. “Humano, demasiado humano”.

Decir que los seres humanos formamos parte del mundo natural podría pasar por una obviedad, pero ciertamente no es trivial. De acuerdo a la clasificación científica, los seres humanos “modernos” (alrededor de 160, 000 años de antigüedad según el record fósil) pertenecen a la especie *Homo sapiens*; a la subespecie *Homo sapiens sapiens*; genero *Homo*; familia *Homínido*; orden *Primate*; clase *Mammalia*; phylum *Chordata* y reino *Animalia*. Ahora, si bien nadie dudaría que las capacidades anatómicas y fisiológicas del organismo humano son capacidades naturales, el estatus “natural” de las capacidades mentales tiende a generar escepticismo entre algunos filósofos y aun científicos, por no hablar de los fundamentalistas religiosos.¹ En esta tesis revisaré uno de los rasgos centrales de nuestra vida mental que ha inspirado la idea de que la mentalidad está más allá de lo meramente natural, a saber, la intencionalidad. Lo haré desde una perspectiva naturalista que se distingue de otras por estar inspirada en la coyuntura científica y filosófica que Darwin propulsó hace ya más de cien años. ¿Qué tiene esta capacidad humana “demasiado humana” (Nietzsche, 1880) de especial?

¹ En 1996 ante la Academia Pontificia de las Ciencias, Juan Pablo II acepto que la teoría de la evolución es “algo más que una hipótesis” y que se aplica exitosamente a dar cuenta del “cuerpo humano”. Sin embargo, aseguró que “las teorías de la evolución que, en virtud de las filosofías que las inspiran, explican al espíritu [mente] o bien como emergiendo de las fuerzas de la materia viva o bien como un simple epifenómeno de esa materia, son incompatibles con la verdad acerca del hombre.”

La intencionalidad es la capacidad de ciertos estados mentales de ser *acerca de* algo más, o de *representar* otra cosa. Los pensamientos, las creencias, los deseos e incluso las percepciones tienen contenido, esto es, representan o significan un estado de cosas. En la segunda mitad del siglo diecinueve, el filósofo y psicólogo Brentano (1874) rehabilitó y caracterizó este concepto: para que entre dos cosas, una creencia y un estado de cosas en el mundo por ejemplo, se dé una relación intencional, es necesario y suficiente que el primer *relatum* de la relación sea algo mental. Más aún, si esta condición se obtiene, es posible que el segundo *relatum* posea un “inexistencia intencional”. Así, es posible pensar, creer, desear y percibir no sólo falsamente, sino también algo inexistente. Puesto que en las relaciones reales o genuinas ambos *relata* deben existir, Brentano concluyó que la relación intencional o de representación era una relación *sui generis* irreduciblemente mental: nada que no fuera mental podría exhibir intencionalidad. Ahora bien, sea o no la intencionalidad la “marca de lo mental” como creía Brentano, lo cierto es que parece haber una parte importante de estados mentales que son intencionales. Parece claro entonces que si hemos de dar cuenta de lo mental debemos ofrecer una explicación de la intencionalidad. ¿Qué tipo de explicación?

Para responder a esta pregunta debemos tener en cuenta dos tipos de hechos relacionados con el problema del lugar de la intencionalidad en el mundo natural: (1) los hechos de “cableado y conexión” y (2) los hechos “interpretativos” (Godfrey-Smith, 2002, 2004, 2006, Sterelny, 2003).

El primer conjunto de hechos refiere tanto a los hechos sobre nuestra organización interna (“cableado” o “hardware”) como a los hechos sobre cómo esa organización registra, refleja o monitorea circunstancias externas (“conexión”). Lo relevante aquí son los mecanismos neuronales que subyacen a la cognición así como las relaciones que estos mecanismos, y las capacidades cognitivas que implementan, entablan con el entorno del organismo en la producción de conductas complejas. Puesto que estos hechos son hechos naturales biológicos, un recuento de su evolución y desarrollo es parte integral de cómo la organización neurocognitiva de los organismos se conecta con el mundo.

El segundo conjunto de hechos corresponden a la llamada psicología popular o de sentido común. Esta psicología ofrece una diversidad de maneras de pensar y hablar sobre fenómenos como las sensaciones, las emociones, las percepciones y los pensamientos. El núcleo de esta

psicología es la psicología intencional: la predicción y explicación de las acciones de los agentes en términos de su concepción del mundo, esto es, de cómo creen que es el mundo, y sus deseos, cómo quieren o desean que su mundo sea. La explicación en términos de creencias / deseos la empleamos para interpretar a otras mentes así como a nuestra propia mente por lo que esta forma de explicación recorre nuestra vida social. En este sentido, las nociones de “contenido”, “significado” o “representación” son empleadas en la práctica cotidiana de interpretar a otros agentes como racionales e intencionales mediante la atribución de estados mentales semánticamente evaluables -creencias, intenciones, deseos, expectativas, etc., a los agentes para explicar y predecir su conducta.²

Ahora bien, este elemento intencional de la psicología cotidiana (2) opera –y de manera muy efectiva-, *independientemente* de la explicación científica de los mecanismos neurocognitivos que subyacen a las distintas capacidades mentales, esto es, independientemente de los hechos “de cableado y conexión con el exterior” (1). Uno de los temas centrales de la filosofía de la psicología es la naturaleza y el estatus de la psicología intencional en relación con una comprensión científica de la cognición. La pregunta es si una psicología científica madura debe incluir, reformular o simplemente eliminar las nociones intencionales y semánticas que operan en psicología popular.

Una motivación para unificar este par de hechos de tal suerte que una psicología científica integre las nociones semánticas e intencionales en un marco naturalista radica en que, más allá de la noción de representación intencional de la psicología humana cotidiana, la noción misma de representación parece jugar un rol ubicuo en las explicaciones de la conducta animal (Millikan, 1984, Dennett, 1995, Godfrey-Smith, 1996, Sterelny, 2003). Los animales son máquinas bioquímicas que involucran sistemas sensoriales y de control complejos seleccionados evolutivamente para registrar confiablemente y responder adaptativamente a elementos y circunstancias en su entorno. Y si bien hay diferencias importantes en las maneras que distintos

² Por ejemplo, X se fue de la fiesta antes que llegara Y porque X *deseaba* evitar toparse con Y porque X *creía* que Y seguía molesto con X pues Y tenía la *expectativa* de que X hubiera actuado de cierta manera en cierta circunstancia.

organismos logran esto, podemos decir que gran parte de nuestros estados cognitivos y de otros organismos tienen el poder de *representar* el mundo para producir acciones consecuentes.³

Ahora bien, para que las nociones intencionales y semánticas de la psicología cotidiana puedan jugar un rol en una explicación científica de nuestras capacidades mentales y la de otros organismos, es necesario explicar dichas nociones en términos no semánticos o no intencionales, en particular, en términos naturalistas, esto es, en términos compatibles con formas de explicación científica. El *desiderata* es explicar qué es representar sin invocar nociones semánticas. La idea es que para poder integrar los hechos interpretativos (2) con los hechos científicos (1) debemos mostrar una coordinación de las relaciones naturales entre estados internos y externos con las relaciones semánticas e intencionales empleadas en la psicología cotidiana. El proyecto es el de usar esta coordinación para *integrar* las propiedades semánticas a propiedades más básicas en el mundo natural.

En las últimas décadas ha habido varios intentos de dar cuenta de lo intencional dentro de un proyecto general de *naturalización* de lo mental. ¿Qué propiedades en el mundo natural pueden exhibir intencionalidad, esto es, representar algo más, incluso algo inexistente? La propuesta que revisaré afirma que lo intencional puede ser explicado en términos de la noción de “función” biológica en el sentido teleológico o teleonómico tal y como es empleado en el contexto de la biología de la evolución. La *teleosemántica*, como se conoce esta propuesta, invoca entonces las propiedades funcionales de los rasgos biológicos para explicar el lugar de la mente, y de la intencionalidad en particular, en el mundo descrito por las ciencias naturales. En esta tesis exploraré el éxito de la propuesta teleosemántica de Millikan (1984, 1993, 2004). El objetivo de esta tesis es realizar una exposición crítica de la teoría teleofuncional y teleosemántica de Millikan y defenderla de una de sus principales objeciones.

Esta tesis consta de tres partes. En el primer capítulo introduzco el problema de la teleología natural y presento una exposición crítica y una defensa de la teoría etiológica

³ Para algunos filósofos (Godfrey-Smith, 1996 y Sterelny, 2003), la comprensión de los elementos distintivos de los sistemas humanos de control requiere integrar las neurociencias y la psicología cognitiva con la biología evolutiva, en particular con la ecología de la conducta humana donde la “integración” depende fundamentalmente de comprender la *función* de estos sistemas. En este sentido, lo que hace posible hablar de “representación” en el contexto de organismos tan diferentes como una abeja, una rana, un castor, una rata o un humano, es que para comprender la operación de los sistemas complejos es necesario comprender su *función*; y para comprender su función es necesario dar cuenta de las relaciones ecológicas, pasadas y presentes, que estos organismos guardan con su entorno.

histórica. En el segundo capítulo expongo una versión teleosemántica y en el tercer capítulo defiendo a la teoría de una objeción central. La estructura del argumento supone que la viabilidad de la teleosemántica depende o tiene como premisa la viabilidad del (teleo)funcionalismo en biología evolucionista. En otras palabras, la naturalización de la intencionalidad en el contexto de la psicología cognitiva depende de una noción naturalista de teleología en el contexto de la biología de la evolución. Podemos presentar el hilo argumentativo de la tesis en tres momentos siguiendo a Walsh (2002):

1. La explicación de la *función biológica evolutiva* en términos de la *historia causal* (etiología).
2. La explicación del *contenido cognitivo* en términos de la *función cognitiva*.
3. La explicación de la *función cognitiva* en términos de la *función biológica evolutiva*.⁴

Tenemos entonces dos retos para el naturalismo filosófico, la teleología natural y la representación mental. Dichos retos hacen patente uno de los puntos de encuentro entre la filosofía de psicología intencional y la filosofía de la biología evolucionista.⁵ A modo de introducción, a continuación presento una breve caracterización del problema de la naturalización de lo mental, del concepto de “intencionalidad”, así como los elementos centrales del enfoque evolucionista con respecto a estos problemas.

2. Naturalismo y Mente.

El programa naturalista en filosofía de la mente es el de *incorporar* lo mental –y en particular lo intencional- en el mundo natural, esto es, en el mundo descrito por las ciencias naturales, tanto las ciencias duras (física y química) como las llamadas “ciencias especiales” (biología y

⁴ En el primer capítulo abordaré el primer punto, mientras que en el segundo capítulo revisaré el segundo y tercer punto.

⁵ Para ello la exposición debe mostrar la convergencia de ciertas nociones y conceptos en común entre las disciplinas, en particular, las que expresan propiedades *relacionales* y, en algunos casos, *normativas* como “contenido intencional”, “verdad” “falsedad”, por un lado, y “función”, “adaptación”, “selección” y “aptitud”, por el otro.

psicología).⁶ La idea central es mostrar que las propiedades mentales y por tanto las propiedades semánticas de los estados mentales son propiedades naturales capaces de jugar un rol en explicaciones científicas de los fenómenos que generan. Si el mundo natural comprende a las propiedades y los estados físicos, químicos, biológicos, geológicos, neurológicos, neurofisiológicos, etcétera, la tarea para el naturalista es la de ofrecer un argumento para incluir en él las propiedades, sucesos y estados mentales-intencionales.

Hay varias motivaciones para una visión naturalista de la intencionalidad. Una de ellas es que las propiedades semánticas no parecen formar parte de las propiedades naturales más básicas como la “energía” o los “quarks”, por lo que las propiedades semánticas o intencionales no deben ser tomadas como primitivas o inanalizables, sino que es necesario dar cuenta de cómo las propiedades semánticas *derivan* de propiedades no-semánticas (Fodor, 1987). Otra razón para no invocar nociones semánticas o intencionales a la hora de dar cuenta naturalistamente de dichas propiedades es la de evitar la circularidad viciosa, esto es, explicar los conceptos semánticos mediante otros conceptos semánticos.

Ahora bien, una de las características de la intencionalidad es la cuestión de la “falsa representación” o el “problema normativo”, el cual supone que los estados a los que atribuimos contenido poseen “condiciones de corrección”, sea en términos de condiciones de verdad o condiciones de satisfacción, y como tal pueden ser correctas o incorrectas. Así, el error cognitivo es posible y la teoría debe dar cuenta de él.⁷ Sin embargo, no se trata sólo de una cuestión sobre

⁶ Para muchos filósofos contemporáneos el naturalismo se identifica con el fisicalismo o materialismo: el naturalismo pretende explicar qué rol, si es que alguno, tienen las propiedades mentales y otros fenómenos *prima facie* no físicos en el mundo físico. Una razón para identificar lo natural con la física es que la física, a diferencia de otras ciencias, parece ser causalmente completa: los efectos físicos siempre tienen causas físicas, lo cual implica que cualquier suceso capaz de producir un efecto físico no puede ser un suplemento ontológico al ámbito de lo físico; de lo contrario habría una proliferación de sobredeterminación causal. Más sobre esto en un momento.

⁷ Una de las principales motivaciones para el escepticismo filosófico frente al estatus “natural” de la intencionalidad, al menos entre los filósofos inspirados en Wittgenstein –*via* Kant, Hegel, los pragmatistas o todos lo anteriores (Sellars, 1956, Kripke, 1984, Davidson, 1970, 1974, 2001a, 2001b, 2001c, 2004, McDowell, 1994, 1998a, 1998b, Brandom, 1994, 2001), es que los estados mentales, en particular las propiedades intencionales de tales estados, son propiedades *normativas* y las propiedades naturales son causales pero no normativas. Estos filósofos defienden una versión de la tesis de la irreducibilidad de Brentano (1874) en términos de la vieja dicotomía “hecho-valor” (Kant, Hume) o “causa-razón” (Wittgenstein) o *questio de jure* y *questio de facto*. Esta dicotomía epistemológica y/o semántica suele presentarse como una superación de la aun más vieja dicotomía ontológica “mental-físico” del dualismo cartesiano (de hecho, la tesis de Brentano puede ser interpretada como una versión del dualismo cartesiano basado en la característica intencionalidad de lo mental), aunque también la dicotomía “mental-físico” puede ser formulada semánticamente en términos de un dualismo conceptual (Davidson, 1974, Sellars, 1956, McDowell, 1994, Brandom, 1994). En estos términos, el naturalismo puede ser pensado como un intento por superar

el error, sino que está estrechamente relacionada con la posibilidad de tener estados representacionales cuyas condiciones de corrección no son satisficibles, por ejemplo, cuando pensamos en algo inexistente como “Dios”, “ninfas” o quizá “la paz perpetua”.

Finalmente, está la cuestión de la *determinación unívoca* del contenido, lo que supone que estados mentales como las “actitudes proposicionales”⁸ son individuadas de forma muy fina tal que dos representaciones extensionalmente equivalente pueden diferir sin embargo en contenido. Por ejemplo, las expresiones “criatura con riñones” y “criatura con corazón” tienen la misma extensión, pues todas las criaturas que tienen riñones tienen corazón y, por lo tanto, ambas expresiones refieren al mismo objeto, pero no tienen el mismo contenido. Este fenómeno es el de la *intensionalidad* (con “s”), donde términos o representaciones son intensionales si y sólo si el que posean el mismo contenido no sólo requiere que tengan la misma extensión sino la misma intensión (Carnap, 1947).

¿Qué alternativas hay dentro del naturalismo para la intencionalidad?⁹ A grandes rasgos, el naturalismo *reduccionista* supone que la naturalización de lo mental y de lo intencional depende de que haya un tipo de reducción de lo mental (propiedades, sucesos, estados o conceptos) a lo no mental (propiedades, sucesos, estados o conceptos), en particular requiere la reducción de lo intencional a lo *no* intencional. El supuesto metafísico de esta versión reduccionista es la de que lo mental tiene una esencia no mental. El supuesto epistemológico es que es posible ofrecer condiciones necesarias y / o suficientes para la reducción de una propiedad o concepto mental a uno no mental, o de una propiedad intencional a una no-intencional. Sin embargo, una dificultad con esta alternativa reduccionista es que suele ser incapaz de distinguirse de las posiciones más radicales conocidas en filosofía de la mente como *eliminacionistas* (Churchland, 1984, Stich, 1994), las cuales, más que naturalizar lo mental, lo eliminan: si una propiedad o concepto mental se reduce a una no mental, entonces en principio es

estas dicotomías. Cabe señalar que la tesis naturalista, sea de corte Darwinista o no, puede ser interpretada tanto en sentido epistemológico como en sentido metafísico (Danto, 1967, Tye, 1992, Sober, 2001). La idea es que la *identificación* de las *propiedades* mentales con propiedades naturales está internamente relacionado con la *explicación* de los *fenómenos* mentales en términos de explicaciones científicas particulares.

⁸ En el siguiente apartado explicaré que son estas “actitudes proposicionales”.

⁹ Hay varios tipos de naturalismo de lo mental que no consideraré aquí. Algunos se distinguen porque tratan de fundarse o bien en una teoría empírica-científica o bien en sobre una base a priori de “análisis conceptual” o en ambos (ver Tye, 1992).

posible reducir un vocabulario mental a uno no mental; y si hay tal reducción, entonces en principio es posible eliminar el vocabulario reducido, esto es, el mental.

El naturalismo *no reduccionista*, en cambio, afirma que lo mental es el objeto de estudio de la psicología científica (cognitiva o fisiológica), y que esto basta para justificar la naturalización de lo mental. Para ello ha de mostrar que la psicología es suficientemente semejante a las ciencias naturales y que lo mental participa en interacciones causales que se rigen por leyes y teorías científicas. Una manera de hacer esto es mostrar que la psicología es semejante a la biología o geografía, química, física, etc. Así, el naturalismo no reduccionista parece tener más un enfoque epistemológico que metafísico y esto no parece ser suficiente para el proyecto de naturalizar lo mental.

Esta insuficiencia está motivada por la idea de que las respectivas ontologías de las distintas ramas científicas están unificadas: vivimos en un mundo físico pero con poderes causales a distintos niveles de realidad.¹⁰ La idea de esta unificación ontológica es que hay cierta tensión entre el proyecto de conceder poderes causales a estados mentales y cierto principio de *clausura causal* en el mundo físico. Según éste, lo físico está causalmente relacionado entre sí y con ningún suceso o estado que no sea físico, ya que todo suceso físico tiene solamente causas físicas y solamente efectos físicos. Además, las relaciones causales se entienden como causas eficientes que se dan espacio-temporalmente y que están subsumidas por leyes. En principio, sin embargo, creemos que los estados mentales están relacionados causalmente entre sí y con sucesos físicos.¹¹ ¿Cómo hacer compatible esto con el principio clausura causal?

Una manera de respetar el principio de clausura causal es mediante una teoría de la identidad que propone que lo mental *es* simplemente lo físico, pero esto supone un reduccionismo eliminativista, el cual –como ya vimos– es indeseable. Otra manera de respetar ese

¹⁰ Por ejemplo, no tiene sentido describir a un organismo en términos de partículas atómicas si lo que queremos es explicar qué tipo de organismo es, lo cual requiere conocer la genealogía histórica de su linaje, cuestión que está indirectamente ligada con la composición microfísica del organismo. Además, el tipo de complejidad organizacional que involucran los sistemas biológicos parece irreducible a meras explicaciones físico-químicas, entre otras cosas, por el rol constitutivo que juega la historia y el entorno así como por las complejas interacciones entre niveles, genético, epigenético, conductual, y simbólico (en el caso humano).

¹¹ Mis creencias contienen información específica sobre sucesos, cosas o estados de cosas extra-mentales y mentales, mis deseos causan acciones que incluyen movimientos corporales; la luz reflejada en los objetos causa mi percepción del objeto; etc. En este sentido, puesto que los hechos mentales frecuentemente tienen efectos físicos en la forma de movimientos corporales (acciones) que intervienen en el mundo extra-mental, estos hechos no pueden estar fuera del ámbito físico. De lo contrario, dichos movimientos tendrían tanto causas físicas como no físicas y no queda claro cómo algo no-físico pueda ser causalmente eficiente del todo.

principio es estableciendo una relación de dependencia ontológica entre las propiedades mentales y las propiedades físicas, una relación de *superveniencia*. Según esta relación, no es posible que haya dos objetos, estados, sucesos o mundos que sean físicamente idénticos y que no lo sean mentalmente, pero puede haber objetos, estados, sucesos o mundos que sean mentalmente idénticos pero físicamente distintos. Esto es posible porque una propiedad mental *M* cualquiera no es idéntica a una propiedad física *F*, sino que puede instanciarse en objetos (estados, sucesos o mundos) que instancian alguna otra propiedad física (*F1*, *F2*, *F3*, ... *F_n*). *M* tiene así una *realización física múltiple*.¹²

Ahora bien, un mundo sin organismos con propiedades mentales no supondría pérdida alguna de una propiedad física causal. El hecho de que los estados mentales sobrevengán en estados físicos no debe excluirlos de ser causalmente eficaces independientemente de sus realizadores. Pero es cierto que la mera relación de superveniencia tampoco garantiza su eficacia causal. En cambio, si las instanciaciones de las propiedades causales tuvieran una relación además de *constitución* con las instanciaciones de las propiedades físicas que son causalmente eficaces, entonces no habría problema en suponer que lo mental es causalmente eficaz, ni se violaría el principio de clausura causal. La noción relevante de constitución se puede obtener mediante la noción de superveniencia *mereológica* (Kim, 1984) - la superveniencia de las propiedades del todo en las propiedades y relaciones de sus partes donde el todo es más que la suma de sus partes. Para mostrar que este tipo de superveniencia es suficientemente robusta para garantizar la eficacia causal de las propiedades supervenientes independientemente de sus realizadores, podemos recurrir a un par de ejemplos.

Pensemos en un choque. Un coche choca de frente con una motocicleta golpeándola con la defensa y ocasionando que el conductor de la moto salga proyectado. Un suceso *C*, el coche

¹² Las propiedades psicológicas o mentales no son el único, ni siquiera el paradigma, de propiedades múltiplemente realizables. En general, las explicaciones que apelan a funciones que son asignadas a ciertas estructuras, mecanismos o entidades en otras ciencias especiales como la biología, la economía o la sociología son múltiplemente realizables. Incluso las propiedades morales, sean o no independientes de las intenciones y propósitos humanos, puede decirse que sobrevienen en el mundo físico. En este sentido, recurrir a la idea de superveniencia requiere explicar el sentido específico en que esta dependencia se logra para que la explicación no sea trivial. Para hacer la tesis de la superveniencia relevante es necesario explicar qué de las propiedades mentales hace que una diferencia mental se deba a una diferencia física, esto es, ofrecer un recuento de la naturaleza de estas propiedades que explique por qué sus instancias deben ser metafísicamente determinadas por hechos físicos. (Sobre el modo en que la teleosemántica satisface esta condición ver el final del apartado 4 de esta introducción y las conclusiones del capítulo segundo).

chocando, causa otro suceso P , la proyección. C es la causa, P el efecto y el choque es el mecanismo que liga la causa con el efecto. Ahora bien, ¿cuál es la propiedad causalmente eficaz que da cuenta de la producción de P a partir de C ? En estricto sentido diríamos que es la defensa del coche, $d(C)$. Sin embargo, puesto que la defensa es una parte del coche y el coche sobreviene sobre sus partes, *pirma facie*, podemos explicar la producción de P sólo mencionando C y por tanto podemos afirmar que C es causalmente eficaz en la producción de P . Otro ejemplo que ilustra la superveniencia mereológica entre propiedades macro y micro es ofrecido por Segal y Sober (1991). Alguien enciende un fósforo en una habitación. Podríamos pensar que el aire dentro de la habitación es causalmente relevante para el encendido del fósforo. Pero alguien podría decir que la propiedad causalmente eficaz del aire es el oxígeno y no el aire puesto que el aire posee otras propiedades, como el nitrógeno, que son causalmente inertes en la producción del encendido. Sin embargo, lo que la superveniencia mereológica muestra es que la propiedad de ser aire y la propiedad de ser oxígeno están a distintos niveles (la primera es una macropropiedad mientras que la segunda es una micropropiedad) y no tenemos que pensar que las propiedades supervenientes al nivel macro tienen que competir con las partes de sus bases supervenientes a nivel micro por la eficacia causal que se obtiene al nivel macro (Segal y Sober, 1991, p. 15). Podemos concluir que las propiedades macro gozan de eficacia causal y por lo tanto que el aire es causalmente eficaz en la producción del encendido.¹³ Así, un estado mental puede

¹³ Segal y Sober (1991, p. 15) recogen esta idea así: Si (i) es una ley (posiblemente no estricta) que todo suceso F causa un suceso G y (ii) en cada caso en el que un suceso F causa un suceso G existen micro-propiedades $m(F)$, $m(F)'$ y $m(G)$ tal que el ser causa de F mereológicamente sobreviene en su ser $m(F)$ y el ser efecto de G mereológicamente sobreviene en su ser $m(G)$ y la posesión de $m(F)$ incluye la posesión de $m(F)'$ y el ser causa de $m(F)'$ causa el ser efecto de $m(G)$, entonces F es eficaz en la producción de G s. La noción requerida de inclusión de la eficacia causal de las propiedades macro es simplemente que la posesión de $m(F)$ es posesión de una conjunción de propiedades uno de cuyos conjuntos es $m(F)'$. Sober y Segal concluyen que las propiedades semánticas de las representaciones mentales son eficaces en las transacciones de representaciones y la producción de la conducta como el aire es eficaz en las transacciones con cerillos. Sin embargo, Segal y Sober (1991) refuerzan la superveniencia mereológica de Kim con el requisito de que la relación causal entre sucesos o propiedades sea nomológicamente simultánea: F mereológicamente sobreviene en $m(F)$ si y sólo si es nomológicamente necesario que si un objeto x tienen $m(F)$ en un tiempo t , entonces x tienen F en t . La idea es que si F mereológicamente sobreviene en $m(F)$, entonces el que los x sean $m(F)$ realiza que los x sean F , es decir, el que los x sean F está fundado en que los x sean $m(F)$.

En oposición a esto, la propuesta teleosemántica que revisaré en esta tesis supone que los estados intencionales no son definidos en términos de leyes, ni siquiera estadísticas, sino en términos de generalizaciones histórico-selectivas (Millikan, 1984, 1986). En este sentido, la relación de superveniencia no sólo no es nomológica sino que no es causal en el sentido sincrónico. La idea es que identificar las propiedades semánticas con propiedades histórico-selectivas no implica que las propiedades intencionales sean causalmente inertes puesto que las propiedades histórica-selectivas no lo son. En cambio, la teleosemántica restituye a los efectos seleccionados

estar compuesto de un estado físico sin tener que ser idéntico a él o alguna parte específica y por tanto la superveniencia se sostiene y al mismo tiempo podemos conceder a la propiedad superveniente eficacia causal, evitando así el epifenomenismo.

El problema es que el carácter intencional de al menos una parte (sino es que todos) los estados mentales no parece *prima facie* permitir tal relación de constitución. La razón es que las propiedades físicas, en particular, las neuronales, son locales mientras que la intencionalidad es algo *relacional* con un entorno externo. Además, lo intencional se caracteriza por ser normativo y ciertamente las propiedades físicas no los son. Tenemos entonces que los estados intencionales deben ser realizados por estados físicos de nivel inferior. Pero dado que la idea de contenido intencional es relacional en tanto involucra al mundo externo, los estados físicos relevantes deben incluir estados externos. ¿Qué estados físicos son “amplios” o “gruesos” en este sentido?

Ésta es una pregunta sobre el tipo de explicaciones científicas que realizan generalizaciones a partir de propiedades físicas amplias, esto es, propiedades dependientes del entorno y la historia. Encontramos estas explicaciones en la biología de la evolución y la ecología. En particular, las adaptaciones y otros rasgos biológicos adaptativos de los organismos son típicamente especificadas en relación a aspectos del entorno (ver nota 35 del primer capítulo). En este sentido, la premisa que afirma que las propiedades físicas son locales es prematura, pues si bien las propiedades descritas por la ciencia física son intrínsecas, podemos perfectamente hablar de propiedades físicas amplias como las que operan en explicaciones biológicas (evolutivas y ecológicas).

3. Contenido y representación mental.

Brentano (1874) caracterizó a la intencionalidad de dos formas diferentes. Por una parte, habló de la característica “dirección a un objeto”, “objeto intencional” o “referencia a un contenido”.

pasados caso de ciertos estados cognitivos una relevancia causal en la producción de casos de esos estados en el presente. En pocas palabras, para el teleosemanticista, la eficacia causal del contenido intencional no es actual o sincrónica, sino que depende de relaciones históricas causales entre los estados intencionales y lo que representan. Así, la teleosemántica evita el epifenomenismo puesto que los efectos con los que identifica el contenido de los estados intencionales son, si bien históricos, causalmente eficaces considerados diacrónicamente en la producción de otros estados intencionales y de conductas específicas. (Más sobre esto en el apartado 4 y en las conclusiones del capítulo segundo).

Se trata de la capacidad de los estados mentales de ser *sobre* o *acerca de* cosas, por ejemplo cuando decimos que el pensamiento de Ana es *sobre* la llegada de la primavera. Por otra parte, hablé de “inexistencia intencional” que se refiere a la aparente peculiaridad de los objetos del pensamiento que dichos objetos pueden ser contenidos del pensamiento aunque de hecho no existan. Esto corresponde a lo que hoy se conoce como el problema de la falsa representación o el problema normativo, esto es, a que los estados intencionales pueden ser sobre lo que no es el caso. Ahora bien, los estados intencionales pueden ser sobre lo que no es el caso en dos sentidos. En el caso de representaciones de cosas que no existen, como cuando pienso en dragones, o en el caso de representaciones falsas, por ejemplo, cuando alguien cree que en México la pobreza es un mito.¹⁴

Independientemente de si ambas caracterizaciones refieren a dos propiedades o son dos caras de la misma moneda (ver Millikan 2004 para una discusión), la noción de “contenido” es una manera genérica de hablar que incluye la noción de “significado” y “representación”. En general, son las oraciones o los actos de habla y los estados mentales intencionales el tipo de cosas que paradigmáticamente tienen contenido.¹⁵ Un tipo de estados intencionales son las “actitudes proposicionales”. Se trata de actitudes que los agentes tienen frente a una proposición que constituye el contenido de la actitud. Distintas actitudes pueden tener el mismo contenido, por ejemplo, creer que el mundo se va acabar, desear que el mundo se va acabar, esperar que el mundo se va acabar, tener la intención de acabar con el mundo, etc; mientras que actitudes del mismo tipo pueden tener distintos contenidos. Las actitudes proposicionales son estados intencionales que atribuimos a agentes u organismos para explicar la conducta de éstos.¹⁶

¹⁴ ¿Qué son los ‘objetos intencionales’? Para Crane (2002), la idea de “objeto intencional” genera un dilema para las teorías de la intencionalidad: o bien los objetos intencionales son entidades existentes ordinarias, o bien son el tipo de entidades que no existen. El problema es que si afirmamos lo primero entonces nos enfrentamos a la dificultad de darle sentido a la idea de que hay objetos intencionales inexistentes como “Pegaso” o “Chupacabras”. Pero si afirmamos lo segundo, entonces estaríamos aceptando la problemática idea de que existen entidades inexistentes. Para Crane la manera de salir de este dilema es rechazando el supuesto de ambas posiciones, a saber, que ser un objeto intencional es ser una *entidad* de cierto tipo. En cambio, la manera adecuada de comprender la noción de Brentano es pensando al objeto intencional como simplemente el objeto (gramatical), para un sujeto, de un acto o estado intencional.

¹⁵ En esta tesis me ocuparé del contenido mental exclusivamente. No hablare del significado lingüístico ni me ocuparé de la cuestión sobre qué tipo de contenido intencional es originario y cuál es derivado, si el lingüístico o el mental. No obstante, la idea misma de atribuir estados representacionales a organismos no humanos, *ergo* no lingüísticos, supone que el contenido original es mental.

¹⁶ Estas propiedades pueden ser interpretadas de forma realista o de forma antirrealista. En el primer caso atribuir una actitud proposicional es atribuir una propiedad que se identifica o sobreviene en un estado mental del

Atribuir una creencia a x es decir que x cree que P . Con ello queremos decir que cierta relación expresada por ‘cree’ se da entre x y el objeto o “contenido” de la creencia denotada por el complemento oracional ‘que P ’. El verbo transitivo ‘cree’ expresa una relación entre un sujeto u organismo denotado por el término sujeto y una proposición denotada por el complemento oracional que es el objeto del verbo, esto es, el contenido de la creencia (Stalnaker, 1999). Por ejemplo, “David cree que las ballenas son mamíferos” dice que David está en una relación de creencia con la proposición *que las ballenas son mamíferos*.

Ahora bien, las actitudes proposicionales no agotan la clase de estados mentales intencionales sino sólo la clase de estados cuyo contenido es *proposicional*.¹⁷ La idea es que si los estados intencionales son individuados por su “objeto intencional” el objeto intencional de las actitudes proposicionales es un “objeto proposicional”, lo cual quiere decir que el estado es individuado mediante sus *condiciones de verdad*.¹⁸ Podemos entonces entender el contenido semántico de una actitud o un pensamiento como aquello que determina qué condiciones en el mundo son relevantes para determinar el valor de verdad del estado, esto es, el conjunto de condiciones de verdad del estado. Por ejemplo, el contenido de creer o esperar que las lluvias llegarán es algo que es verdadero o será verdadero si y sólo si las lluvias llegan. Así, el que una creencia sea verdadera bajo ciertas condiciones y otra creencia sea falsa bajo esas mismas condiciones es suficiente para que las creencias tengan contenido diferente.¹⁹

Podemos también representar el contenido proposicional de una actitud como el conjunto de mundos posibles en los que la actitud es verdadera (Lewis, 1970, Stalnaker, 1999, Jackson,

agente y éste existe independientemente del acto de atribución. En el segundo caso la propiedad sobreviene en la atribución misma, en cuyo caso tener una actitud es algo que es posible sólo en relación a otro agente, un interprete. No argumentaré a favor del realismo intencional contra la postura interpretativista. Daré por supuesto que los estados intencionales son estructuras internas (neuronales) de algún tipo.

¹⁷ Las emociones por ejemplo, son estados intencionales cuyo contenido no es una proposición: si x ama a y x no ama una proposición. Mientras que otros estados mentales presumiblemente no tienen contenido, como el dolor en donde el elemento esencial es la *cualidad* de la experiencia o se caracterizan porque la ausencia de objeto, como es el caso de acuerdo a ciertas teorías psicológica con la angustia o la ansiedad.

¹⁸ Sin embargo, hablar de estados intencionales no proposicionales como teniendo “condiciones de verdad” es inofensivo y me permitiré este uso para hablar del contenido de representaciones de organismos no humanos. La idea es que este modo de hablar expresa el hecho de que el contenido de una representación expresa la relación entre ésta y ciertas condiciones *en el mundo*. Lo mismo al hablar de “valores de verdad” para el caso de estados cognitivos no proposicionales. En este sentido, hablar de “condiciones de verdad” no pretende implicar ninguna posición realista en el sentido opuesto a una semántica verificacionista o de “condiciones de aceptabilidad”.

¹⁹ Podemos simplemente individuar las proposiciones en términos de sus condiciones de verdad y a las actitudes en términos de su objeto proposicional sin ocuparnos del tipo de estructura que tienen las proposiciones.

2001). La idea es que las creencias y pensamientos que son verdaderos bajo las mismas condiciones son verdaderas bajo los mismos mundos posibles. La intuición es que la noción de “contenido” se puede construir en términos de una división entre posibilidades independientemente del estatus ontológico de estos mundos. Así, representar cómo son las cosas es hacer una división entre lo que concuerda con cómo las cosas son representadas como siendo y lo que no concuerda con cómo las cosas son representadas como siendo. La idea es que si el contenido es la propiedad que determina las condiciones que se deben obtener en el mundo para que la actitud sea verdadera, podemos pensar en esta manera de representar el mundo como determinando ciertas posibilidades o circunstancias –las que hacen a la representación verdadera–, y excluyendo otras, las que la hacen falsa.

Es importante distinguir el *contenido* de una representación y el *vehículo* del contenido, esto es, la representación misma. Pensemos en la imagen de un perro dibujada, en la palabra escrita “perro” y en la palabra dicha “dog”. Lo que tenemos son tres representaciones diferentes con el mismo contenido. En el primer caso el vehículo es visual, en el segundo y tercer caso el vehículo es lingüístico donde si bien ambas palabras tienen el mismo valor semántico el vehículo sintáctico es diferente además de que la presentación del vehículo en el primera caso es tipográfica mientras que en el segundo es fonética. En el contexto de la psicología cognitiva y la filosofía de la mente interesada en psicología de sentido común, se ha introducido el término teórico de “representación mental” (Cummins, 1984, Stich y Warfield, 1994, Crane, 1995). Brevemente, éstas son estados cognitivos internos a los sujetos que involucran tres elementos:

- (a) Un *vehículo* de representación, el cual puede ser entendido como un suceso neuronal;
- (b) Un *objeto* u objetivo [*target*] de la representación, que en general es un aspecto del entorno (pero también puede ser otra representación); y

- (c) Un *contenido* de la representación, esto es, lo que hace a un (a) particular del tipo (a) ser una representación particular del tipo (b).²⁰

En este contexto, la intencionalidad hace entonces referencia a la propiedades semánticas de las representaciones mentales, esto es, al hecho de que tales estados poseen (c) en virtud de mantener una *relación* con algo, un objeto u estado de cosas, en el mundo. ¿Qué clase de relación es ésta? La intencionalidad de (c) puede ser caracterizada con base a tres características:

- (i) Es una relación entre un estado, signo o representación y otra cosa que ella misma tal que el primer *relatum* es “sobre de”, “acerca de”, “dirigida a” o “significa qué” -en una palabra, *representa* el segundo *relatum*.²¹
- (ii) Un estado, signo o representación intencional puede ser acerca de lo que no es el caso. Esta *normatividad* corresponde a la posibilidad de error o falsedad que emerge del carácter semánticamente evaluable de los estados intencionales, esto es, de que a un estado intencional le correspondan condiciones de corrección (sean condiciones de satisfacción o condiciones de verdad).
- (iii) Un estado, signo o representación intencional puede representar algo que no existe (representación vacía), es decir, el segundo *relatum* de la relación puede ser inexistente.²² La idea es que las

²⁰ En esta tesis me ocuparé fundamentalmente de (c) aunque su relación con (a) y (b) es constitutiva de la idea misma de contenido. Por ejemplo, podemos localizar la posibilidad de que un particular de tipo (a) sea falsa si hay una falta de correspondencia entre (c) y (b).

²¹ Por ejemplo, mi estado mental de desear una tomar unas vacaciones es un caso de relación intencional en la que el objeto intencional es algo extra mental: algo “mental”, mi deseo, se relaciona con algo extra mental (un estado de cosas en el mundo), el tomar unas vacaciones. Sin embargo, el contenido de un estado mental intencional puede perfectamente ser otro estado mental (intencional o no), como cuando pienso acerca de otra creencia, por ejemplo, cuando pienso que mi creencia de que la religión es falsa y nociva para el hombre es una creencia popular o no. Aquí mi creencia tiene por objeto otra creencia. También podría pensar *acerca* de otros estados mentales no intencionales, como por ejemplo el dolor que sentí ayer, etc. Finalmente también hay representaciones que son acerca de ellas mismas o auto-referentes como en el caso de la paradoja del mentiroso.

²² Ésta es la paradoja de Brentano. Puesto que en las relaciones reales o genuinas (relaciones físicas naturales) ambos *relata* deben existir, y puesto que la relación intencional o de representación es tal que uno de los *relatum* puede no existir, la intencionalidad es una relación *sui generis*, irreduciblemente mental. Más aún, es la

atribuciones de estados intencionales resisten la cuantificación. (Esta es la propiedad que enfatizó Brentano para demarcar lo mental de lo físico.)

En este sentido, una teoría naturalista del contenido semántico de las representaciones mentales se enfrenta al problema fundacional de la (psico)semántica: ¿Cómo puede algo representar otra cosa? O más exactamente, ¿cómo es posible que una relación real, natural exhiba las características (1), (2) y (3)? Para circunscribir el problema que me ocupa en este ensayo y explicitar el sentido de “fundacional” conviene distinguir entre distintos problemas relacionadas con la intencionalidad (Stalnaker, 1999):

1. El problema *fundacional* de explicar *filosóficamente* la intencionalidad.
2. El problema *empírico* de explicar los mecanismos *psicológicos* de representación mental.
3. El problema *técnico* de explicar *semánticamente* la estructura de los enunciados de atribución de estados intencionales.

La idea central detrás del naturalismo es que (1) y (2) deben ser compatibles. Para ello es necesario que (1), en algún sentido, explique las propiedades semánticas en términos de propiedades naturales. Una forma de hacerlo es reduciendo las propiedades semánticas a alguna clase de propiedades naturales de suerte que la condiciones de identidad de las primeras puedan darse en términos de alguna clase de las segundas, ya sea en el sentido de identidad de tipos o en el sentido de que las propiedades semánticas sobrevengan en las naturales. La idea es que (1) debe ser caracterizado apelando a propiedades naturales de tal suerte que la noción de “representación” en (2) sea científicamente adecuada. Si bien la respuesta a (2) es en último término una cuestión empírica, la viabilidad de (2) depende de la resolución del problema

marca de lo mental frente a lo físico. Sobra decir que éste es uno de los factores por los cuales el fenómeno de la intencionalidad ofrece tantas resistencias a un tratamiento naturalista.

fundacional que supone (1).²³ En otras palabras, dado el problema fundacional que aqueja a la idea misma de representación, no está claro qué contaría como respuesta a (2) a menos que tengamos una versión adecuada de qué es *representar*.²⁴ En este ensayo no me ocuparé de (3) sino fundamentalmente de (1) y de la manera en que las respuesta a (1) arroja luz sobre (2). Sin embargo, es necesario decir algo sobre (3) para motivar la necesidad de atender el problema de los fundamentos de la (psico)semántica independientemente de otros problemas relacionados.

El problema semántico técnico (3) es el de explicar cómo el significado o contenido de un enunciado de atribución de una creencia esta determinado como una función del significado de sus partes constitutivas. La estructura semántica básica de tales enunciados es de la forma *x cree que P* o *[C (x, p)]*. En términos generales, podemos decir que dicho enunciado afirma que una relación expresada por ‘cree’ se da entre el individuo denotado por el término del sujeto ‘x’ y aquello que es denotado por la cláusula oracional ‘que P’. El problema es entonces especificar qué tipo de cosa es denotada por la cláusula oracional y cómo dicha cosa es determinada como una función de los constituyentes de la cláusula.²⁵ En este sentido, el problema semántico es el problema de la estructura composicional de proposiciones (u oraciones) que ocurren en oraciones de atribución. Para los fines de ese tipo de análisis, ‘cree’ simplemente expresa una

²³ Éste es un caso de la relevancia de la filosofía para las ciencias: atender un problema conceptual que, si bien no es una cuestión meramente *a priori*, tampoco puede resolverse sólo empíricamente.

²⁴ Otra opción naturalista que mencioné en el apartado anterior es la idea de que quizá una psicología intencional no tiene cabida dentro de una psicología (neuro)cognitiva científica. En cuyo caso, la noción de intencionalidad que subyace a la idea de “representaciones mentales” debe ser *eliminada*, por lo que es posible que los mecanismos cognitivos relevantes (ii) no resulten ser representacionales en el sentido de la intencionalidad *sensu* Brentano.

²⁵ Hay distintas teorías sobre lo que los objetos de las creencias son. Un problema con individuar el objeto proposicional en términos de “condiciones de verdad” o de “mundos posibles” es que nos compromete con una manera distinguir contenido que no da cuenta de la fina individuación que se sigue de los hechos relativos a las atribuciones de los estados intencionales como las creencias. Estos hechos sobre la atribución de creencias sugieren que los objetos de las creencias deben ser individuados de forma más fina que en términos de condiciones de verdad: dada dos cláusulas oracionales diferentes como ‘que P’ y ‘que Q’ es posible encontrar un contexto en el que es plausible decir que alguien cree que P pero no cree que Q incluso si es el caso que P si y sólo si Q. Si P y Q son necesariamente equivalentes (o “intensionalmente isomórficas” o sinónimas), entonces los ejemplos muestran que la equivalencia necesaria, el “isomorfismo intensional” o la sinonimia, no es suficientes para la identidad de las proposiciones. Sin embargo, hay razones para resistir extraer una concepción de las condiciones de identidad de las proposiciones u objetos de las creencias a partir de la opacidad característica de los contextos de atribución de estados intencionales (Stalnaker, 1999). Un recuento teórico que explique las creencias y otros estados cognitivos intencionales en términos de capacidades y disposiciones a interactuar con el mundo extra-lingüístico motivan en cambio una individuación más gruesa del contenido de las creencias en términos de condiciones de verdad. En este sentido, podemos aceptar que dos estados tienen el mismo contenido (dos creencias por ejemplo les corresponde el mismo objeto proposicional) si son verdaderos bajo las mismas condiciones.

relación entre el creyente y el objeto de la creencia. Así, la solución al problema semántico (3) se puede dar independientemente de dar cuenta de la naturaleza misma de esta relación (1). En cambio, el problema fundacional es precisamente el de caracterizar dicha relación, y en concreto, hacerlo en términos naturalistas.²⁶

4. Teleosemántica.

El naturalismo pretende explicar las propiedades semánticas de las representaciones mentales en términos de propiedades naturales. Lo que distingue a la teleosemántica es que las propiedades naturales que invoca para explicar las propiedades semánticas de los estados mentales son propiedades funcionales, en el sentido teleológico empleado por la biología evolucionista: la idea es que las condiciones de verdad de estados cognitivos como las creencias y las condiciones de satisfacción de los estados conativos como los deseos se explican en términos de las funciones biológicas de dichos estados.²⁷

²⁶ En el capítulo tercero veremos cómo la objeción de la indeterminación (Fodor, 1990) a la propuesta teleosemántica en parte surge de confundir (i) con (iii).

²⁷ La teleosemántica se inscribe en el contexto más amplio de las aproximaciones evolucionistas y Darwinistas a la psicología en general y a la cognición en particular, tanto animal como humana. En este sentido, la teleosemántica no está exenta de supuestos empíricos. Concretamente, supone la hipótesis (histórica) de que la cognición en general, y la mente humana en particular, son *adaptaciones* complejas al medio ambiente (ver nota 5, capítulo primero). Ahora, esto no implica que la cognición humana se originó por selección natural. Bien podría tener un origen exaptativo. Pero sí implica al menos que la selección natural ha mantenido y modelado el complejo de rasgos en cuestión. No entraré aquí al debate sobre el estatus empírico de las hipótesis adaptacionistas, en particular a las que conciernen a la cognición humana con los característicos problemas metodológicos que se siguen de la falta evidencia disponible. Asumiré sin ninguna pena –como el propio Darwin–, que la cognición está conformada por varias adaptaciones, y en particular que la característica plasticidad adaptativa de la cognición es una de ellas (ver Sterelny, 2003, Neander, 1999, Sterelny y Griffiths, 1998, Millikan, 1993, Dennett, 1995, Godfrey-Smith, 1996 entre otros). La intuición general que habla a favor de este supuesto es la siguiente: suele decirse que el cerebro humano involucra el más alto grado de complejidad organizacional en el universo. Sea o no esto el caso, lo cierto es que la selección natural es la única fuerza en la naturaleza, o si se prefiere, la única forma de explicación capaz de dar cuenta de este grado de complejidad, no sólo en el mero sentido de *orden* sino el sentido robusto de *diseño*. Además existe otro problema que debe ser mencionado, a saber, el que compete a la cuestión de la *individuación* del rasgo que hemos llamado “mente” o “cognición” (ver Godfrey-Smith, 1996). Esto hace factible que la cognición no sea reducible a un solo rasgo y por tanto que carezca de una sola historia evolutiva capaz de incluir las distintas capacidades que asociamos con la mentalidad y con la cognición en particular. De ahí que haya hablado arriba en plural de “complejo de rasgos” y de “varias adaptaciones”. Ahora bien, independientemente de cuántos rasgos subyacen a la cognición, la cuestión que me compete es una propiedad relativamente general de la cognición *simpliciter*, a saber, su carácter representacional. En este sentido, la hipótesis es que los rasgos cognitivos, independientemente de cuantos sean y cuáles, son individuados como todo rasgo que sea una adaptación, esto es, relativamente al entorno externo y a la historia.

La noción relevante de función biológica es teleológica, no causal o disposicional. La función de un órgano no se identifica con una causa (eficiente) actual o presente del objeto de la atribución funcional sino con algún *efecto por el que fue seleccionado*. El proceso relevante que subyace a la atribución funcional es entonces el proceso de selección natural. Este proceso requiere (a) variación es un proceso mecánico causal que opera sobre variaciones genotípicas y fenotípicas de los rasgos relevantes que son el resultado de mutación genética aleatoria y no de un proceso deliberado. La tesis es que este proceso “ciego” es sensible a la funcionalidad de los rasgos que selecciona.

En términos generales, la propuesta teleosemántica de Millikan (1984) es que las relaciones donde uno de los *relatum* no existe en el sentido no ser actual, no es una propiedad irreduciblemente mental, sino que la exhiben las estructuras o rasgos biológicos evolucionados que tienen funciones, esto es, las adaptaciones. Por ejemplo, la función propia del corazón es bombear sangre, la de los riñones es la de limpiar la sangre, y un rasgo biológico puede tener un función (o varias) aunque no la ejecute o aunque la ejecute defectuosamente. Así, un corazón que no pueda bombear sangre o uno riñones que limpie la sangre defectuosamente aun así retienen sus respectivas funciones. En otras palabras, el propósito biológico de un rasgo, como las condiciones de satisfacción de una representación intencional, puede no cumplirse. Puesto que lo que hace a una representación ser intencional es que puede ser falsa, esto es, que puede representar lo que no es el caso, la apuesta teleosemántica es que la *falsa* representación es un caso de *disfunción*. Así, la normatividad relevante que permite evaluar a las representaciones mentales sin apelar a nociones semánticas es una normatividad que puede ser construida en términos de normas biológicas históricas.²⁸

Ahora bien, la teleofuncionalidad si bien da cuenta de una propiedad de la intencionalidad, ser sobre lo que *no* es el caso, no exhibe *ipso facto* la capacidad de ser “acerca de” algo. Claramente ni el corazón ni los riñones son “acerca de” algo; no representan nada en el sentido intencional. Hace falta decir qué tipo de funciones propias caracterizan a los rasgos

²⁸ Lo que tenemos entonces es que la fuente de la teleofuncionalidad es la selección natural y la fuente de la intencionalidad es la teleofuncionalidad. Dicho en otras palabras, las “representaciones mentales” derivan su contenido de un proceso selectivo (teleonómico) no intencional. Así, la teleosemántica promete matar dos pájaros de un tiro: (1) la normatividad de las representaciones mentales deriva de sus funciones biológicas y (2) la fuente de dicha normatividad es la “normatividad natural” que encontramos en las funciones biológicas, esto es, que pueden o no llevar acabo su “propósito” por el que fueron seleccionados y, sin embargo, tenerlo.

cognitivos y cómo son ejecutadas dichas funciones (qué mecanismos son relevantes para ello). Esto me ocupara en el tercer apartado del cuarto capítulo.

Un problema sustancial para una psicología cognitiva madura que pretenda incluir a la intencionalidad dentro de su repertorio de propiedades “naturales” es si se debe caracterizar a sus sujetos de forma tal que se haga referencia a cómo están integrados en su medio ambiente, esto es, mediante una propiedad relacional o extrínseca, o si simplemente se debe hacer referencia a lo que sobreviene en las estructuras de los cuerpos y los cerebros de los individuos u organismos, esto es, mediante una propiedades locales. Una consecuencia central del enfoque teleonómico de la cognición es el papel que el medio ambiente y la historia juegan en la caracterización tanto de lo que la cognición *es* como del papel o *función* que tiene. Esta consecuencia, trasladada al contexto de la intencionalidad implica una clara postura anti-individualista o externista con respecto a la individuación del contenido de las representaciones mentales. En este sentido, el que una creencia corresponda con el mundo externo como “se supone que debería hacerlo” de acuerdo a las funciones de los mecanismos que las producen o en cambio sea falsa, es una cuestión relativa a la manera como encajan el organismo y su medio. Ahora bien, uno de los propósitos de este ensayo es dar una caracterización positiva de este “anti-individualismo” en el sentido preciso de un “externismo ecológico-evolutivo”, partiendo de la premisa de que la inclusión del medio ambiente y la historia son elementos constitutivos de la caracterización que demos no sólo de las conductas intencionales de los organismos sino de sus capacidades representacionales que presumiblemente subyacen a tales conductas.

Una ventaja de esta propuesta en el contexto de la filosofía naturalista de la mente es que recurrir a propiedades funcionales otorga un contenido preciso a la tesis de la superveniencia. Las propiedades funcionales son paradigmáticamente propiedades que no son reducibles –en el sentido de tipos- a propiedades o hechos físicos. Por ejemplo, ninguna propiedad física es necesaria y suficiente para ser un ojo o un ala. Para ser un ojo o una ala se necesita, por un lado una estructura física adecuada y por otro tener la función de ver o de volar, de tal suerte que hay una gran variedad física de cosas que pueden realizar esta función tal como se hace patente en la historia evolutiva de estos rasgos. Esta condición se cumple para otras categorías funcionales en biología, tanto para rasgos fisiológicos y anatómicos como para rasgos a nivel genético o cognitivo-conductual. Así, independientemente de que no se obtenga una reducción tipo entre

propiedades funcionales y propiedades o hechos físicos, es indiscutible que las propiedades funcionales están metafísicamente determinados por propiedades físicas. La idea es que dos cosas no podrían ser idénticas físicamente sin ser idénticas funcionalmente. Ahora, es importante calificar esta tesis. Para que dos cosas físicamente idénticas sean funcionalmente idénticas las propiedades físicas en cuestión no hacen referencia sólo a las propiedades intrínsecas micro-estructurales de estas cosas, sino a las propiedades físicas amplias que incluyen tanto al entorno como a la historia causal. Luego, dos cosas pueden ser idénticas en términos de sus propiedades físicas intrínsecas y, sin embargo, diferir en sus funciones si no comparten las mismas propiedades históricas y ecológicas.

Finalmente, una característica central de la propuesta teleosemántica es que las funciones biológicas son *efectos* y no directamente causas. Más bien son efectos *pasados* que a su vez causan la presencia de tales efectos en el presente en virtud de que las estructuras que producen dichos efectos se reproducen en virtud de tales efectos. La idea es que los efectos juegan un papel causalmente eficiente, considerados históricamente. Así, la naturalización del contenido semántico no se realiza identificando al contenido con alguna propiedad causal (disposicional) actual de un sistema cognitivo, sino con ciertos efectos que el sistema posee en virtud de su historia causal. Podemos concluir que explicar las propiedades semánticas de las representaciones mentales en términos de propiedades funcionales implica que las propiedades semánticas no son idénticas a las propiedades físicas y, sin embargo, explica porque los hechos semánticos sobrevienen en hechos físicos, naturalizando así al contenido intencional. El programa teleosemántico no reduce las propiedades semánticas a propiedades físicas sino a propiedades biológicas, en particular, a propiedades evolutivas y ecológicas. En este sentido, la estrategia naturalista teleosemántica está muy lejos de eliminar las propiedades de orden superior a favor de propiedades de orden inferior.

CAPÍTULO PRIMERO:

Teleología Natural, Funciones Biológicas y Teoría Etiológica

“(...) it is difficult for anyone who tries to make out the use of a structure to avoid the word purpose”

Charles Darwin

1. Introducción: teleología natural.

En el contexto de la filosofía de la biología y la filosofía de la psicología, podemos hablar actualmente de dos enfoques generales sobre funciones biológicas, la propuesta teleológica, iniciada por Larry Wright (1973) y la del rol causal propuesta por Robert Cummins (1975). Por su relevancia para el problema de la representación mental (capítulo segundo), revisaré principalmente la propuesta teleológica o “etiológica” y versiones más sofisticadas de ella. Pero también presentaré la propuesta del rol causal. Si bien tanto las funciones teleológicas como las del rol causal tienen aplicaciones tanto en biología como en psicología y sociología, en este capítulo me limitaré al contexto biológico. La estructura del capítulo es la siguiente: en el apartado segundo expondré la forma de la explicación teleológica y la propuesta (proto) etiológica de Wright (1973). En la tercera parte presentaré la teoría etiológica histórica de Millikan (1984), enfatizando el carácter normativo de las “funciones propias” así como el modo en que esta teoría resuelve los contraejemplos que afectan a la propuesta de Wright. En la cuarta parte presentaré la principal alternativa a la teoría etiológica, la teoría propensionista de Bigelow y Pargetter (1987). Defenderé la propuesta etiológica de sus objeciones. En la quinta parte expondré el énfasis “modernista” de Godfrey-Smith (1994) a la propuesta histórica de Millikan. En la sexta y última parte presentaré las funciones de rol causal (Cummins, 1975) así como las posibles relaciones entre este tipo de funciones y las etiológicas defendiendo un pluralismo desunificador. En el resto de esta sección me ocuparé de presentar el problema de la teleología natural y la pertinencia de dicha noción en el contexto de la biología de la evolución.

La pertinencia de la teleología en el mundo biológico puede hacerse patente con un ejemplo (Hull, 1974). Pensemos en dos fenómenos naturales, uno físico como calentar un gas y

otro biológico como calentar a un mamífero. Mientras un físico diría que calentar un gas causa que se expanda, un biólogo diría que calentar a un mamífero causa que éste sude. Hasta aquí no hay problema; tenemos una causa previa que produce un efecto posterior. Sin embargo, el biólogo podría además decir que un mamífero suda cuando es expuesto al calor *para* conservar su temperatura constante (termorregulación), mientras que ningún físico diría que un gas se expande cuando es calentado “para” mantener su temperatura constante, incluso cuando eso es exactamente lo que ocurre.

¿Qué quiere decir aquí el “para”? Una intuición obvia sobre la diferencia entre el sistema físico y el sistema biológico es pensar en el efecto de sudar causado por calentamiento como teniendo el “propósito” o “meta” de enfriar la temperatura del cuerpo del mamífero. En oposición diríamos que la expansión del gas es un efecto “ciego” o simplemente “mecánico” del calentamiento. En este sentido, sudar es un proceso *teleológico*, mientras la expansión del gas es un proceso mecánico no-teleológico. Ahora bien, cuando atribuimos un “propósito” a algo debemos preguntarnos por el tipo de proceso que le subyace a la atribución y si este proceso puede fundar una atribución teleológica.

Podemos decir que el proceso mecánico de calentar un gas no es sensible a la teleología precisamente porque no tiene sentido –como sí lo tuvo en la física aristotélica- decir que el gas tiene el propósito de expandirse: el fenómeno físico puede ser exhaustivamente explicado mediante causalidad eficiente sin invocar un proceso que involucre algún tipo de propósito.

Donde tiene perfecto sentido atribuir propósito es en el contexto mental. Si mi propósito de matar a alguien es el de vengarme, el proceso en el que la teleología de mi acción se apoya es un proceso psicológico, en particular, un proceso intencional: realizo la acción correspondiente de asesinar movido por la intención de vengarme. La acción involucra, de una forma que ni la expansión del gas ni la acción de sudar involucran, la *representación* explícita de las condiciones de realización de aquello que constituye el objeto del propósito. Lo que subyace al propósito en este caso es un agente intencional: un sistema cognitivo y conativo que actúa *para* lograr la

realización de su propósito (o deseo, o intención) guiado *por* la representación de éste, de suerte que su acción es *intencional*.²⁹

Ahora, si bien está clara la pertinencia de la teleología en el contexto mental y su irrelevancia en el contexto físico, ¿qué ocurre en el contexto biológico? Hemos dicho que intuitivamente los procesos teleológicos suelen presuponer intencionalidad, esto es, la representación explícita del “propósito” o “meta” particular, tal que dicha representación antecede a su realización. Sin embargo, es obvio que no podemos presuponer intencionalidad o alguna otra propiedad psicológica en el contexto biológico. La pregunta es entonces, ¿qué proceso natural (biológico) podría fundar las atribuciones teleológicas sin recurrir a un proceso psicológico intencional?³⁰

La teoría etiológica está inspirada por un intento de explicar el uso de expresiones teleológicas en el contexto de la biología de la evolución. El núcleo de la propuesta es que el “para”, en el ejemplo del mamífero, no apela a un agente intencional sino al proceso mecánico histórico de evolución por selección natural: los mamíferos presentan el rasgo fenotípico de sudar, porque dicho efecto incrementó la aptitud de tales organismos en el pasado frente a otras variaciones del rasgo, fijando así el rasgo a través de las generaciones en la población. Podemos decir que el propósito de sudar que es el de regular la temperatura por ese efecto causó la selección de dicho rasgo. Si bien “propósito” en este contexto no causa nada, pues no preexiste a su efecto, la idea es que hablar de propósito en este sentido sí implica su realización causal eficiente por el tipo de proceso que le subyace o que lo realiza, a saber, la selección natural. Veamos otro ejemplo.

¿Cómo provee la selección natural una explicación de *por qué* los humanos poseen un corazón en virtud de *para qué* son los corazones? El corazón es una estructura con diferentes capacidades o disposiciones causales que pueden generar distintos efectos dadas ciertas

²⁹ En este caso, parece que el propósito es *causa* de la acción al mismo tiempo que determina el *objetivo* de la misma.

³⁰ En otras palabras, si lo que queremos es una descripción *naturalista* de la teleología natural, bien haríamos en ofrecer una forma de explicación capaz de *reducir* la teleología natural a la noción o a alguna de las nociones –presumiblemente la evolutiva–, de “función biológica”, empleadas en la práctica y en el discurso biológico. En cualquier caso, la idea central de este apartado es que la biología evolucionista emplea la noción de “función” de forma teleológica en ciertas explicaciones –paradigmáticamente, en las explicaciones *adaptacionistas*–, de una manera que parece no enteramente reducible a los términos de la causación familiar de las ciencias físicas.

condiciones (que pueden ser ambientales o internas al organismo). Por ejemplo, producir el particular sonido de su latido o el contribuir al color característico de la sangre, bombear sangre, etc. Pero si queremos explicar cómo es que los seres humanos llegaron a tener corazones, ciertos efectos resaltarían en su importancia o utilidad evolutiva frente a otros efectos. Así, en contraste con estos efectos accidentales, bombear sangre parece ser la disposición o el efecto característico de dicha estructura que nos permite identificarlo como su función. Presumiblemente, bombear sangre es el efecto *por* el que hubo selección de la variante actual de corazones en la población humana.³¹ La idea central es que la selección natural explica la presencia de un rasgo explicando cómo *fue* o *es* preservado después de haber sido producido de forma aleatoria, a causa de ciertos efectos particulares que aumentan la aptitud relativa del organismo al medio.

2. Explicación teleológica y teoría (proto)etiológica.

¿En qué sentido apelar a la “teleología”, “propósito” o “función” de algo tiene poder explicativo? El *explanandum* de una explicación teleológica es la existencia o presencia del objeto de la atribución funcional. La explicación teleológica pretende responder a la pregunta ‘¿Por qué está ahí?’ respondiendo a la pregunta anterior de ‘¿Para qué es eso?’. Decir que la función de un rasgo como sudar es mantener constante la temperatura del cuerpo explica *por qué* los mamíferos sudan, esto es, por qué poseen dicho rasgo. Larry Wright (1973) caracterizó la forma general de la explicación teleológica así: decir que la función de X es Y significa que:

³¹ El corazón es un órgano complejo y los rasgos complejos suelen ser el producto de modificaciones acumuladas en estructuras antecedentes a través de su historia filogenética. Estas modificaciones a estructuras o rasgos existentes ocurren de manera aleatoria como resultado de mutación genética o recombinación del material genético. Cuando esto ocurre encontramos variación en la población con respecto a dicho rasgo fenotípico. Si alguna de las variantes del rasgo provee a sus poseedores con una ventaja en la lucha por la supervivencia y la reproducción, entonces esa variante será mejor representada en el genotipo de la población de las siguientes generaciones debido al efecto fenotípico de bombear sangre. Cuando esto ocurre, los biólogos evolucionistas afirman que esa variante del rasgo fenotípico ha incrementado la aptitud relativa de su poseedor al bombear sangre. Por supuesto, esa variante constituye a su vez la base para nuevas modificaciones, tanto en la estructura como en la función de dicho rasgo. En este sentido, los humanos tienen corazón porque los corazones fueron el producto de modificaciones generadas de manera aleatoria sobre estructuras preexistentes que fueron preservadas o mantenidas por medio de la selección natural a causa de que proveían a sus poseedores con una ventaja competitiva (Sober, 1984, pp. 147-155), en particular, debido a que bombeaban sangre y lo hacían mejor que otras variaciones de corazón con la misma función antes de que se fijara el tipo de corazón moderno en la población de los humanos.

X existe porque realiza o tiene el efecto Y, y

Y es una consecuencia o resultado de que X exista

La virtud del modelo etiológico es que, en lugar de entender a Y como una causa de X, el análisis parte del hecho de que Y es un *efecto* de X y toma a este hecho como una de las causas eficientes *antecedentes* de X. La explicación etiológica es una explicación mediante causas *antecedentes*. Una estructura o entidad puede tener o producir distintos efectos según sus capacidades causales. Algunos de estos efectos son accidentales mientras que otros parecen constitutivos de la estructura. La explicación de la presencia o persistencia de una estructura en un sistema se realiza citando aquellos efectos que no son meros accidentes. En este sentido, asignar una función a algo es atribuir una propiedad que permite discriminar los efectos que son funciones de los que son meros efectos accidentales.

¿Cuál es el criterio que nos permite identificar los efectos accidentales de los que constituyen la función o funciones de un rasgo? Wright no da cuenta de la fuente de dicha distinción en términos de algún proceso específico, sino que lo deja abierto en su formulación: el proceso es relativo al contexto relevante de explicación en la que se realiza la atribución funcional. Sin embargo, si bien dejar abierto a consideraciones pragmáticas el proceso que subyace a la identificación de ciertos efectos de una entidad con sus funciones gana en generalidad, éste tiene el inconveniente de generar contraejemplos como los siguientes (Boorse, 1976):

(A) Considérese a una roca pequeña sosteniendo una roca más grande en una corriente rápida en movimiento. Si la roca pequeña no soportara a la grande, sería aplastada por ella. Sostener a la roca grande es lo que la roca pequeña hace y eso explica por qué está ahí. Bajo la formulación de Wright, ésta es la función de la roca pequeña.

(B) Considérese ahora una manguera diseñada para transportar monóxido de carbono al exterior. La manguera accidentalmente desarrolla una lágrima y el gas que se

fuga causa que el mecánico quede inconsciente e incapaz de reparar la fuga. Bajo la propuesta de Wright, la lágrima tiene la función de mantener al mecánico inconsciente.

3. Teoría etiológica histórica.

Millikan (1984) llama “funciones propias” a las funciones teleológicas y propone que lo que cuenta como la función propia de un rasgo está determinado por la historia evolutiva de dicho rasgo, en particular, por su historia selectiva: la función o funciones de un rasgo son los efectos que causaron su selección en el pasado y que explican la presencia de dicho rasgo así como del efecto relevante en el presente. Esta versión evita los contraejemplos arriba señalados, al ofrecer dos restricciones a la formulación original de Wright:

1. Atribuir una función a un objeto individual es algo que se realiza relativamente a una *clase* a la que el objeto pertenece, donde dicha “clase” es *histórica*. Para explicar la pertenencia a una clase histórica, Millikan (1984) introduce el término técnico de “familia-reproductiva-establecida” (FRE). La idea central de FRE es que sus miembros deben estar ligados entre sí causalmente como *ancestros* y *descendientes* para ser miembros de la misma clase histórica, del mismo “linaje”. Hay dos tipos de FRE:
 - (a) Un individuo es un miembro de una FRE de *primer orden* si tiene algunos rasgos en común con otros miembros de esa familia resultado de un proceso de reproducción, en donde “reproducción” es entendido en el sentido de un proceso de copia gobernado por leyes que causan el mantenimiento, a través de las generaciones, del rasgo o rasgos en cuestión.
 - (b) Un individuo es un miembro de una FRE de *orden superior* si tiene algunos rasgos en común con otros miembros de la familia como resultado de pertenecer a miembros de la FRE de primer orden. Esto quiere decir que los miembros de

una FRE de orden superior no son resultado de la copia o reproducción directa sino resultado de la reproducción o copia de miembros de una FRE de primer orden.

2. Poseer un efecto funcional es algo que *beneficia* al objeto o sistema contenedor. La ejecución de la función de un rasgo causa una diferencia en la *aptitud* del organismo portador a su entorno. A diferencia de Wright, la distinción entre “efecto funcional” y “efecto accidental” está fundada sobre la base de un proceso causal *específico*, a saber, la selección natural. Los “efectos funcionales” se distinguen de los “accidentales” por ser efectos *seleccionados*, esto es, efectos que causaron una diferencia en la *aptitud* del organismo portador (aumentaron su probabilidad de sobrevivir y reproducirse al hacer Y) y que se transmiten hereditariamente.

La idea central de la teoría etiológica histórica es que **la función o las funciones de un rasgo es el efecto o son los efectos *por los que dicho rasgo fue seleccionado***. Sober (1984) introdujo la distinción entre “selección de” y “selección por” para dar cuenta de la estructura causal de la teoría neo-darwinista. Brevemente, la “selección *de*” refiere a los *efectos* del proceso de selección la “selección *por*” describe las *causas* del proceso de selección. La idea es que hay selección *de* ciertas propiedades neutrales de los objetos o rasgos seleccionados, por ejemplo, al hacer ruido del corazón, pero *por* propiedades o efectos adaptativos, en este caso, bombear sangre. Lo que esta distinción muestra es que la selección natural sólo puede operar sobre contribuciones causales actuales *pasadas* a la aptitud. Así, el efecto o efectos por el que la estructura o rasgo biológico *fue* seleccionada y que explica la presencia de dicho efecto seleccionado en el presente, se vuelve *retrospectivamente* la función de la estructura. En resumen, el que la historia sea “evolutiva” implica la transmisión de una generación a otra del rasgo funcionalmente caracterizado. Por otra parte, el que la historia sea “selectiva” implica que el rasgo resulta *útil*, esto es, que beneficia al organismo en términos de supervivencia y reproducción, en virtud de su función. Así, “heredabilidad” y “utilidad”, explican la presencia y el mantenimiento de un rasgo funcional en una población.

Ahora podemos regresar a la formulación de Wright a la luz de las restricciones de Millikan. ¿En qué sentido hacer *Y* o tener el efecto *Y* de *X* explica la presencia de *X* dadas las restricciones (1) y (2)? La función de *X* es hacer *Y* sólo en el caso que hacer *Y* causó la proliferación de los ancestros que poseían *X*. Lo relevante a notar aquí es que no es el efecto *Y* en el presente de la entidad funcional *X* en el presente, sino una entidad similar *pasada*, un ancestro de *X*, el que es responsable de causar la existencia de la entidad en el presente, al hacer aquello que llamamos su función, esto es, *Y*. A la base de la teoría etiológica histórica está la idea de que **los efectos *pasados* tipo proveen la explicación de la existencia de casos *actuales* de ese tipo.**³² Ahora podemos dar la *definición* teórica de “función propia”:

Definición (df): un caso actual de un rasgo *X* de un organismo *O* (miembro de una *FRE*) en un entorno *E* tiene la función de producir un efecto del tipo *Y*, si y sólo si, casos pasados (recientes) de *X* contribuyeron a la aptitud de los ancestros de *O* en *E* al producir *Y* y fueron seleccionados *por* hacer esta contribución a la aptitud de los ancestros de *O* frente a variaciones *X* en *E*.³³

Millikan (1984) llama a estas explicaciones históricas que respaldan la atribución funcional explicaciones “Normales”. Una explicación Normal es una explicación de cómo una *FRE* particular ha históricamente ejecutado una función propia particular. Esta explicación puede ser más o menos próxima en el tiempo. Por ejemplo, puede ser una explicación sobre el origen o el episodio original de selección, o sobre episodios recientes de mantenimiento de dicho rasgo

³² La selección es siempre de *tipos* (de genotipos y sus expresiones fenotípicas) no de *casos*. Las atribuciones funcionales, consecuentemente, tienen por objeto los tipos y sólo de forma derivada los casos.

³³ Las propiedades funcionales o “efectos seleccionados” de los rasgos de los organismos son propiedades “amplias” o “gruesas” y se atribuyen a rasgos conocidos en biología evolucionista como “adaptaciones”. “Adaptación” es un concepto ecológico, relacional. La adaptabilidad es una relación entre la forma orgánica o el rasgo fenotípico y el entorno. Esta adaptabilidad de los rasgos se explica por selección natural. Dada cierta variación heredable de un rasgo adaptativo tipo, la selección natural favorece a las variaciones que produzcan un incremento en la aptitud del organismo portador, esto es, en su capacidad de sobrevivir y reproducirse. Cuando un rasgo es mantenido dentro de una población por selección natural, llamamos a dicho rasgo una adaptación. La propiedad por la que fue seleccionado el rasgo, su efecto adaptativo, es lo que se llama la función del rasgo. Podríamos preguntar, ¿porqué un organismo *O* tiene el rasgo *X*? Si *X* es una adaptación, la respuesta es que *X* existe porque produce el efecto *Y*, y hacer *Y* incrementó la aptitud de los ancestros de *O* en el entorno ecológico *E*, causando la selección de *X* por *Y*. La *función* de un rasgo (una adaptación) es una propiedad o estado “grueso” o “amplio”, no basta la *forma* del rasgo sino su *relación* con el entorno es necesaria. Más no suficiente. La *utilidad*, el beneficio de esa relación para el organismo, determina su naturaleza. Las adaptaciones son rasgos que son individuados a partir su funciones, de los efectos por los que fueron seleccionados. Para la teleosemántica, como veremos en el siguiente capítulo, éste mismo patrón individua las representaciones mentales y por lo tanto constituye su contenido.

frente a variaciones en competencia. Las “condiciones Normales”, por su parte, son las condiciones ambientales a las que una explicación Normal ha de hacer referencia. Éstas son las condiciones bajo las cuales se produjo el episodio de selección relevante. La explicación Normal responde a las preguntas históricas: ¿bajo qué condiciones fueron seleccionados los ancestros de *X* por hacer *Y*?, y/o también ¿cuáles eran las condiciones del entorno ecológico tal que hacer *Y* beneficiaba al organismo portador de *X* y tal que hacer *Y* causó la selección de *X*? Millikan habla de “Normal” con mayúscula para enfatizar que dichas condiciones históricas son normativas, a diferencia de hablar de condiciones “normales” en el sentido de promedio. La idea es que no se trata de que la mayoría de las veces *X* ejecutó *Y*, sino que aunque *X* haya ejecutado *Y* la minoría de las veces, tales ocasiones son las decisivas en términos de aptitud. El carácter normativo de las funciones propias se expresa en el hecho de que si *Y* es la función *X*, entonces *Y* es lo que *se supone que X debería hacer* en condiciones Normales.³⁴

A su vez, la distinción normativa fundamental es entre ejecución propia o correcta de una función y ejecución disfuncional o defectuosa (incorrecta) de una función. La idea es que un rasgo biológico funciona propia o correctamente cuando hace aquello por lo que fue seleccionado y para hacer esto es necesario que las condiciones de ejecución sean históricamente normales. Así, en ausencia de condiciones normales “históricamente óptimas”, la ejecución de la función es defectuosa o disfuncional. Es importante recalcar que el tipo de normatividad al que apela Millikan es una normatividad histórico evolutiva, esto es, un sentido de normatividad que se debe distinguir de una normatividad evaluativa o estadística:

- No es evaluativo porque juzgar que algo funciona propiamente o ejecuta su función correctamente no es lo mismo que juzgar que su funcionamiento es bueno. Neander (1995) pone el ejemplo de la fertilidad de los adolescentes. Biológicamente es normal; sin embargo, de esto no se sigue que la fertilidad en los adolescentes sea algo bueno o deseable, sino quizá lo contrario.

³⁴ En adelante utilizaré “normal” o “normales” para referirme a lo que Millikan llama “Normal” y “Normales”, excepto cuando un contraste entre este tipo de condiciones históricas y condiciones promedio o estadísticas sea pertinente.

- Por otra parte no es una normatividad estadística, como lo muestran las enfermedades pandémicas y las epidemias. Por ejemplo, si todos los corazones de todos los mamíferos dejaran de bombear sangre, aun así sería la función propia de este rasgo bombear sangre. Otro ejemplo es el de los espermatozoides (Millikan, 1984). Una de las funciones propias de los espermatozoides es la de fecundar al óvulo (otra función propia más próxima es la de nadar en dirección al óvulo). Sin embargo, sólo uno de entre millones logra de hecho ejecutar su función propia de acuerdo a una explicación normal. En este sentido, la mayoría estadística de los espermatozoides nunca ejecutan su función y, no obstante, la poseen, y aun así, al parecer, suficientes espermatozoides realizan su función como para que dicha entidad siga difundiéndose como miembro de una FRE.

Ahora bien, ¿qué consecuencias tiene que la normatividad sea histórica? Principalmente, que para Millikan las normas son hechos naturales, en particular hechos históricos que permiten discriminar los hechos actuales que concuerdan o no con los hechos pasados. En este sentido, la *función propia* atribuida a un rasgo y las *ejecuciones actuales* de dicha función no están al mismo nivel: no todas las circunstancias *actuales* en las que una función putativa se ejecuta son condiciones *históricamente ideales* o normales. La idea de Millikan es que podemos discriminar las condiciones “actuales” y las condiciones “ideales” (“históricamente optimas”) sólo si conocemos la historia evolutiva particular de la entidad en cuestión, esto es, recurriendo a una explicación normal. La consecuencia importante es que la propiedad normativa funcional, si bien es una propiedad *real* y natural de la estructura o entidad biológica funcional, al ser una propiedad histórica, no propiedad *actual* de dicha estructura, ni microestructural ni disposicional.

En relación al naturalismo, lo que destaca es que esta normalidad que surge del conocimiento de la historia selectiva del rasgo funcionalmente caracterizado, permite hablar de una “normatividad natural”: atribuir una función, esto es, fijar un efecto o efectos seleccionados de entre efectos accidentales, es establecer un criterio que permite discriminar ejecuciones correctas e incorrectas de la función, esto es, un criterio normativo con respecto al cual las ejecuciones de la función pueden concordar o no. El criterio normativo que se sigue de atribuir

una función propia es que la ejecución correcta es la que corresponde a condiciones normales y a una explicación normal correlativa, mientras una ejecución disfuncional corresponde a condiciones anormales donde tales condiciones son condiciones *del entorno*, tanto interno como externo al organismo. Podemos entonces hablar genuinamente de normas biológicas evolutivas.³⁵

La idea central de Millikan es que las entidades biológicas, como el corazón o los espermatozoides (pero también las capacidades cognitivas representacionales) al ser objeto de atribución funcional, no caen dentro de esta categoría en virtud de su constitución actual o de sus disposiciones, sino en virtud de su historia, en particular, en virtud del patrón etiológico relevante dado por su historia causal selectiva. El concepto biológico de “función evolutiva” en el sentido de la teoría etiológica histórica, describe propiedades *relacionales* y *normativas* entre el organismo y el medio que *benefician* al organismo cuando son ejecutadas normalmente.

4. Teoría Propensionista.

Para Bigelow y Pargetter (1987), en cambio, describir la función de un rasgo biológico es describir una estructura presente en términos de un suceso o estado de cosas *futuro*: la función de los riñones en tiempo t es limpiar la sangre en el tiempo t' donde $t' > t$. Puesto que las estructuras funcionalmente caracterizadas no necesariamente producen los efectos funcionales, los sucesos o efectos futuros relevantes en muchos casos son inexistentes; muchos riñones fracasan en limpiar la sangre. Ahora bien, un efecto o evento futuro *inexistente* no constituye una propiedad causal genuina de un rasgo biológico. Por lo tanto, las propiedades funcionales no pueden tener un rol en explicaciones científicas-naturales de la presencia del rasgo funcionalmente descrito. Puesto que la presencia del rasgo y de sus propiedades actuales no depende causalmente de la función de éste, el rasgo y sus propiedades hubieran existido incluso si no hubiera dado lugar a los efectos o sucesos que constituyen su función. Bigelow y Pargetter concluyen que las explicaciones funcionales son *redundantes*. Éste problema afecta el núcleo de la teoría etiológica: no podemos explicar la evolución y la presencia de un rasgo en una

³⁵ Por ejemplo, el corazón existe por el efecto seleccionado de bombear sangre, es decir, por dicho efecto es que el corazón se hace presente en cada nueva generación. Sin embargo, puede haber corazones defectuosos o condiciones externas anormales que impidan bombear sangre adecuadamente. En dicho caso decimos que el corazón tiene la función de bombear sangre, pero que la ejecuta incorrecta o disfuncionalmente, o simplemente no la ejecuta.

población afirmando que dicho rasgo ha evolucionado y esta presente *por que* sirve una función específica.

El principal defecto de la teoría etiológica, para Bigelow y Pargetter, es que la dirección de explicación es hacia el pasado en lugar de hacia el futuro, y cómo tal, la teoría etiológica es incapaz de generar predicciones como demanda toda “genuina” explicación científica. Más aun, éste defecto tiene la consecuencia de hacer a la teoría etiológica vacua:³⁶ podemos juzgar que un rasgo tiene cierta función sólo *retrospectivamente*, esto es, después de que el efecto relevante ha de hecho contribuido a la aptitud del organismo [y ha sido seleccionado].³⁷ Más aun, el que un efecto sea parte de la función de un rasgo esta constituido por el hecho de que tener dicho efecto ya ha contribuido a la aptitud del organismo con ese rasgo [y a que dicho incremento de la aptitud ya ha causado la selección del rasgo]. La teoría etiológica explica entonces qué efectos son funciones *post factum*, así, no explica porqué –la causa, de que tales efectos y no otros sean la función del rasgo. La circularidad viciosa de esta forma de explicación radica en que si la función *F* de un rasgo *R* explica la persistencia de *R* en la población *P*, entonces no podemos explicar que *R* tenga *F* recurriendo al hecho de que *R* ha persistido en *P*.

³⁶ El análisis propensionista del concepto de “función” se basa en el precedente del análisis propensionista de la noción evolutiva de “aptitud” (Mills y Beatty, 1979). La propuesta de Bigelow y Pargetter es que lo que se cumple para la aptitud se cumple para las funciones biológicas. En cada caso, el análisis propensionista surge como alternativa a la supuesta vacuidad explicativa de los análisis históricos de éstas nociones (aptitud y función). En el caso de la noción Darwinista de “supervivencia del más apto” la objeción, equivocada pero ilustrativa, es que es tautológica, y por tanto, que la noción de aptitud carece de todo peso explicativo. La idea es que este defecto es producto de que la teoría juzga la aptitud de un organismo sólo en términos *retrospectivos*: sabemos qué organismos fueron los más aptos *post factum*, esto es, después de que hemos visto qué organismos de hecho han sobrevivido. Más aun, la teoría asume que el hecho de que ciertos organismos hayan sobrevivido y otros no es lo que constituye que sean más aptos. Así, la explicación Darwinista de la aptitud no es sólo *post factum* sino *post hoc*. La alternativa propensionista es que la aptitud no debe ser definida retrospectivamente en términos de supervivencia *actual* si no en términos de una propiedad disposicional del individuo en un entorno que da al individuo cierto “potencial de supervivencia” o ventaja reproductiva. Podemos entender a la propiedad disposicional como una propiedad subjuntiva: dicha propiedad especifica qué pasaría o qué podría pasar si se obtienen ciertas condiciones o circunstancias que disparan la disposición. Esta teoría evita la putativa circularidad de la teoría Darwinista puesto que la aptitud, al ser una propiedad disposicional (subjuntiva), es una propiedad que sobreviene sobre los rasgos morfológicos del individuo (o especie). Así, la explicación opera indicando que el organismo (o especie) tiene ciertas propiedades casuales activas que, bajo ciertas circunstancias o condiciones, producirán el fenómeno o efecto a explicar: en condiciones apropiadas la aptitud explica la ventaja *actual* en la capacidad de sobrevivir y reproducirse de la misma manera que bajo las condiciones apropiadas la fragilidad de una copa de cristal explica su ruptura *actual*. Tenemos entonces que ‘ser más apto’ no es sólo una consecuencia de que unos organismos hayan sobrevivido y otros no, sino que es la causa de que esto haya ocurrido.

³⁷ En su formulación de la teoría etiológica, Bigelow y Pargetter omiten frecuentemente una referencia explícita a la selección que en el texto pongo entre corchetes. El problema es que el contenido de los corchetes es el elemento central de la teoría etiológica. Más adelante haré explícito los problemas para la teoría propensionista y la crítica que hace de la teoría etiológica derivados de hacer esta omisión.

Bigelow y Pargetter presentan una teoría alternativa que pretende restituir un genuino rol explicativo a las funciones en el contexto biológico capaz de capturar la intuición de que las funciones pueden ser caracterizadas en referencia a sucesos futuros posiblemente inexistentes. La idea central es que la función de un rasgo es una propiedad *disposicional* (subjuntiva) que confiere al organismo portador una *propensión* a incrementar su aptitud bajo condiciones apropiadas que disparan la disposición. Así, cuando las funciones dan lugar a la sobrevivencia, esto es, cuando las condiciones que disparan esta disposición se obtienen, al igual que cuando una disposición se manifiesta, la causa es la forma estructural morfológica del rasgo y la relación de esta forma con el entorno. De la misma manera que las disposiciones sobrevienen sobre sus bases categoriales, las funciones sobrevienen sobre la forma estructural de la morfología del rasgo. Esto hace posible que las funciones sean causalmente eficaces y que puedan jugar un rol genuino en explicaciones científico-naturales, pues al igual que las disposiciones explican sus manifestaciones, las funciones explican la sobrevivencia.³⁸ Podemos concluir que el hecho de que cierto efecto haya contribuido a la aptitud no es una mera consecuencia de que dicho efecto sea parte de la función del rasgo sino la causa de ello.³⁹

¿Qué puede responder el teórico de la etiología? Una diferencia crucial entre ambas teorías es que para la teoría propensionista un rasgo que acaba de aparecer y que produce propensiones a incrementar la aptitud *ya* posee una función.⁴⁰ Ahora, una consecuencia

³⁸ Los autores señalan (p. 109) que la noción de “propensión a incrementar la aptitud” debe ser desarrollada en términos formales con el rigor del cálculo de probabilidad. Así, la idea será determinar precisamente el sentido de “incrementar la *probabilidad* de sobrevivir y reproducirse”. En este sentido, la propuesta de Bigelow y Pargetter no promueve una teoría propensionista particular, sino que promueve la idea de que un esquema propensionista general es el que mejor puede dar cuenta de las funciones biológicas.

³⁹ Es importante destacar que Bigelow y Pargetter motivan la pertinencia de recurrir a “propensiones” a partir del esquema etiológico: El rasgo en cuestión debió haber tenido el efecto relevante en un *suficiente número de ocasiones*, donde estas condiciones son “apropiadas” (condiciones bajo las cuales el efecto contribuye positivamente a la aptitud). Así, la historia que confiere la función a un rasgo exhibe cierto patrón: el efecto relevante que eventualmente devendrá en función debe haber estado ocurriendo las suficientes veces y en los contextos apropiados en al menos un proporción *estadísticamente* significativa de casos (p. 107). Esta contribución a la aptitud en dichos casos no se puede deber, de acuerdo a las leyes de la probabilidad, a un mero accidente. En cambio, dichos casos son casos en los que el rasgo confiere una propensión “sostenida” a incrementar la aptitud del organismo. La idea es que un rasgo tiene una función si ha sido mantenido en virtud de poseer una propensión a incrementar la aptitud. En cambio, si un rasgo es mantenido debido a una secuencia de accidentes ese rasgo no tiene una función. Lo que confiere el estatus de función a un rasgo no es el mero hecho de que la aptitud del organismo incremente en virtud de tener ese rasgo, sino el incremento en la aptitud en virtud de las propensiones que el rasgo confiere al organismo (p. 107).

⁴⁰ La teoría etiológica describe a un rasgo actual como sirviendo a una función cuando en el pasado dicho rasgo “confirió propensiones” que incrementaron la aptitud del organismo portador. Bigelow y Pargetter (p. 108) en

fundamental de esto y que no esta explícita en la presentación de Bigelow y Pargetter, es que esto implica que un rasgo posee una función *antes de e independientemente de que haya selección de ese rasgo por el efecto relevante*. Así que, o bien Bigelow y Pargetter asimilan la noción de selección a la de aptitud, en cuyo caso la estructura conceptual de la teoría de la evolución queda desdibujada (ver Sober, 1984), o bien excluyen a la selección para concentrarse en la aptitud, en cuyo caso su presentación de la teoría etiológica es inadecuada. En cualquier caso esto lleva a importantes confusiones que se hacen patentes en la manera en que Bigelow y Pargetter “derivan” su propuesta a partir de la propuesta etiológica (ver nota 41). A continuación señalaré cuatro consecuencias que se siguen de ésta asimilación conceptual y que pretenden motivar la necesidad de asignar un rol central a la selección natural, y con ello a la perspectiva histórica, en una teoría de la funciones biológicamente adecuada. Posteriormente presentaré la respuesta de Millikan a la crítica de vacuidad explicativa.

(1) Los conceptos de “aptitud”, “función” (adaptación) y “selección” están ciertamente interrelacionadas y constituyen la estructura conceptual de la teoría de la evolución. Sin embargo, no sólo son lógicamente independientes sino que juegan distintos roles dentro de la teoría de la evolución (ver Sober, 1984). De hecho, puesto que la noción de función depende radicalmente de la noción de selección de acuerdo a la teoría de evolución por selección natural, para que una teoría propensionista se aplique a la noción de función tal y como es empleada el biología evolucionista, no basta que se aplique exitosamente a la noción de la aptitud, es necesario ofrecer una teoría propensionista (exitosa) de la selección.⁴¹ Ahora, existen problemas con presentar la selección natural en términos de propensiones, en especial debido a la estructura histórica de las explicaciones evolutivas, pero aunque fuera posible generar una interpretación propensionista de la selección natural, la “propensión a incrementar la aptitud” y “propensión a ser seleccionado” no son lo mismo. El primer tipo de propensión puede producir el segundo tipo si existe variación

cambio proponen que el rasgo ha estado sirviendo esa función desde el principio, esto es, no sólo antes de “ahora” sino antes de que dichas propensiones hayan *de hecho contribuido* positivamente a la aptitud, pues antes de que esto fuera el caso el rasgo confería un propensión a incrementar la aptitud en el organismo (p. 108). Y es esta propensión la que constituye la función del rasgo.

⁴¹ Bigelow y Pargetter, después de identificar la función de un rasgo con la propensión a incrementar las probabilidades de sobrevivir y reproducirse (aptitud) hablan de “propensión a ser seleccionado” (p. 110) e identifican la función de un rasgo con la propensión a ser seleccionado en virtud de que dicho rasgo tiene los efectos relevantes. Sin embargo, no presentan una teoría propensionista de la selección.

en la población, pero un rasgo puede tener el primer tipo sin de hecho ser seleccionado.⁴² Un rasgo puede incrementar la aptitud en virtud de producir un efecto adaptativo para el organismo, pero sino hay variación heredable en la población, dicho rasgo no puede ser objeto de selección. La distinción relevante aquí es entre rasgos con efectos adaptativos y rasgos (adaptaciones) con efectos seleccionados. Para ser lo primero basta poseer un efecto que incremente la aptitud, para ser lo segundo es necesario ser seleccionado por dicho efecto, lo cual presupone variación heredable de dicho rasgo. Para la teoría propensionista, son los efectos conducentes a un incremento en la aptitud los que son la función de un rasgo y no propiamente lo efectos seleccionados.⁴³ El problema con la propuesta propensionista es entonces que analiza la noción de función a imagen y semejanza de la noción de aptitud, pero si bien la noción de aptitud es parte integral de la noción etiológica (evolutiva) de función, la noción fundamental es la de selección.

(2) Una consecuencia fundamental de que la selección determine la función de un rasgo es que éste “proceso pasado” es sensible a los efectos que afectan positivamente la aptitud *independientemente* de si dichos efectos constituyen una muestra significativa en términos estadísticos (ver nota 41). Para la teoría etiológica el “suficiente número de ocasiones” en las que el efecto relevante que retrospectivamente constituye la función se debió haber obtenido *no* tiene que ser estadísticamente relevante, es decir, las suficientes veces no tienen que ser la mayoría de las veces. De hecho, un efecto puede conferir un alto grado de aptitud a un organismo y ser consecuentemente seleccionado, aun si de hecho se obtiene o se obtuvo la minoría de las veces. Por esta razón Millikan (1984) llama a las condiciones “apropiadas” condiciones históricamente “Normales” precisamente para distinguirlas de condiciones promedio (“normales”). Ahora, si las condiciones “suficientes” pueden y frecuentemente son la minoría, como afirma Millikan, entonces *no* podemos derivar una propensión “sostenida” a incrementar la aptitud a partir de efectos que pueden ocurrir la minoría de las veces. En pocas palabras, la noción de propensión depende de que los efectos que confieren un incremento en la aptitud se obtenga un número de

⁴² En esta línea, Neander (1991b) niega que una disposición a incrementar la supervivencia y una disposición a ser seleccionado sea lo mismo. Algunas funciones reproductivas en ocasiones ponen en peligro la vida -por ejemplo, dar a luz, o golpearse el pecho entre los gorilas machos- por lo tanto no incrementan la supervivencia; y sin embargo sí incrementan la aptitud y, por lo tanto, son aptas a ser seleccionadas.

⁴³ Así, contrario a las recomendaciones de Gould y Vrba (1982), para Bigelow y Pargetter las “exaptaciones” tendrían funciones y no sólo efectos (ver nota 35).

veces estadísticamente significativo. Pero puesto que la relevancia estadística es irrelevante desde el punto de vista de la selección, la noción de “propensión sostenida” no puede jugar el papel que Bigelow y Pargetter le asignan.

(3) Ellos afirman que un rasgo tiene una función si ha sido mantenido en virtud de poseer una propensión a incrementar la aptitud. Pero desde la perspectiva etiológica, un rasgo posee una función si es mantenido o desechado por selección natural. Si bien el incremento en la aptitud causa la selección, sin la selección la capacidad de heredar el rasgo en cuestión que implica la idea de selección, la mera aptitud es causalmente inerte para mantener un rasgo. Así, es la selección natural y no la *propensión* a incrementar la aptitud la que mantiene a los rasgos funcionales. Esto es central puesto que la propensión es una propiedad que sobreviene *en* el rasgo dado ciertas condiciones en el entorno, mientras que la selección natural es una fuerza *exterior* al rasgo.⁴⁴ Así, los rasgos no se mantienen o desechan a sí mismos, sino que es una fuerza externa la que lo hace a causa de ciertas propiedades de los rasgos, de ciertos efectos. De nuevo, su “derivación” equivoca el proceso histórico relevante.

(4) Otro problema para el propensionista es que la naturalización del elemento teleológico implícito en las atribuciones de funciones biológicas no se realiza mediante la noción de aptitud sino mediante la noción de selección natural. En particular, el elemento teleológico de éstas atribuciones radica en el “por” de la noción de “selección por” (Sober, 1984). Por lo tanto, una noción de función biológica que no se funde en la noción de selección por, no retiene el elemento teleológico que tanto apela a nuestras intuiciones pre-teóricas.⁴⁵ En esta línea Neander (1991b)

⁴⁴ Claro que las funciones de un rasgo, al igual que sus disposiciones, sobrevienen sobre el rasgo (al final de este apartado hablare sobre estas diferencias para resistir la asimilación de las funciones a disposiciones). No obstante, una diferencia fundamental es que las funciones se fijan por medio de un mecanismo exterior al rasgo, la selección natural. La idea es que para la teoría etiológica decir que un rasgo es mantenido en virtud de la función que sirve implica que la selección mantiene al rasgo por el efecto que constituye su función. El elemento a destacar es que si la propiedad funcional no fuera sensible a la selección no tendría sentido decir que la función explica el mantenimiento del rasgo.

⁴⁵ Para Bigelow y Pargetter el elemento teleológico de nuestras intuiciones pre-teóricas da pie a la idea de que describir una función es describir algo presente a partir de un efecto *futuro*, sin embargo, dichas intuiciones pre-teóricas suponen intencionalidad, esto es, la *representación* del suceso futuro. Así, el precio por conservar el elemento teleológico sin retener la intencionalidad implícita en la noción intuitiva de teleología *-contra* Bigelow y Pargetter, es invertir la dirección de la explicación del futuro al *pasado*. ¿Qué justifica esta inversión? El paradigma de algo que tiene una función en el sentido teleológico es un paradigma psicológico: X mató a Y *para* robarle o con el *propósito* o la *intención* de robarle. El *telos* de la acción no proviene de la acción misma sino de las intenciones y las metas del agente que realiza la acción y que causan la acción correspondiente. Puesto que un recuento naturalista, biológicamente adecuado de lo que es una función teleológica no puede presuponer intencionalidad, la intuición

también responde a la teoría propensionista. Ella acusa a dicha teoría de no capturar el patrón general de las explicaciones teleológicas en la medida en que éstas son genuinas explicaciones causales elípticas.⁴⁶ Las explicaciones teleológicas explican entidades presentes ya que apelar a la función o funciones de dichas entidades involucra una referencia implícita a causas anteriores y la teoría propensionista evita toda referencia a causas anteriores. En el contexto biológico, dichas causas anteriores refieren a sucesos pasados de selección. Así, la teoría de Bigelow y Pargetter tiene el inconveniente de que la selección futura sólo puede explicar instancias futuras de rasgos mas no instancias presentes. En aras de hacer de la explicación funcional una genuina explicación científica -predictiva, Bigelow y Pargetter restringen excesivamente el rol explicativo que tiene atribuir funciones.

Por su parte, para Millikan (1989) dos cosas justifican ver para atrás en lugar de hacia el futuro, (i) la estructura temporal *histórica* de las explicaciones selectivas y, (ii), la indeterminación del valor de verdad de los contrafácticos que se derivan de la formulación propensionista. En cuanto a (i) la idea central es que en una explicación lo que *asegura* la ocurrencia de un suceso no es necesariamente lo que *explica* la ocurrencia del suceso (p.116). Lo que asegura responde la estructura *lógica*, sin embargo lo que explica hace referencia a la estructura *temporal*. La pregunta es, ¿cuál es la estructura temporal relevante de las explicaciones funcionales que apelan a la evolución por selección natural? Para Millikan dicha estructura es intrínsecamente *histórica*, pues apelan causas temporalmente anteriores. En este sentido, ninguna explicación que apele a la selección natural puede “ver hacia delante” como pretenden Bigelow y Pargetter.⁴⁷

sobre el futuro es abandonada a favor del pasado. ¿Qué proceso natural puede anticipar ciertos resultados tal que podamos llamar a tal proceso como genuinamente teleológico? Sólo los procesos psicológicos pueden anticipar en este sentido, pues tales procesos involucran intencionalidad. Ahora, la selección natural es único proceso natural que sin poder anticipar resultados podemos decir que es sensible a diferencias en los efectos de ciertas rasgos tal que podemos llamar “funciones” a dichos efectos una vez que se ha producido selección *por* dichos efectos. En este sentido, en lugar de anticipar, la selección natural retrospectivamente discrimina ciertos efectos de otros.

⁴⁶ “Elíptico” en este contexto hace referencia a la idea de que si bien el propósito o la función no son directa o estrictamente causalmente eficaces, sí implican un proceso causalmente eficaz que subyace a la realización del propósito: la selección natural. En este sentido, las explicaciones teleológicas en biología sí son explicaciones causales. En términos de Millikan, diríamos que se trata de una explicación histórica, esto es, que apele a la historia causal, a las causas eficientes antecedentes y no a las causas eficientes actuales.

⁴⁷ Para Millikan, la confusión que motiva dicha propuesta tiene que ver con distinciones tipo-caso: ciertamente los efectos ocurren después de sus causas, por lo que cada caso individual de función ocurre después de la génesis del rasgo caso que efectúa dicha función. Ahora, el elemento central de las explicaciones selectivas es que efectos pasados de causas pasadas dan cuenta de causas caso actuales –por ejemplo, sobrevivencias pasadas

En cuanto a (ii), la propuesta propensionista supone que un rasgo incrementa o incrementaría la aptitud si, en promedio sobre los individuos actuales en la especie, tener dicho rasgo produciría un individuo más apto que si no tuviera el rasgo (Millikan, 1989). Esta explicación apela a contrafácticos.⁴⁸ Millikan afirma que en este contexto el recurso a contrafácticos es problemático pues su valor de verdad está indeterminado. La razón es que el valor de aptitud de un rasgo es relativo a variaciones alternativa *actuales* de dicho rasgo en la población. Así, es irrelevante desde el punto de vista biológico plantear qué aptitud global tendría un organismo si no tuviera el rasgo pertinente en *ausencia* de variaciones de ese rasgo, e incluso en relación a variaciones del rasgo meramente *posibles*. Ahora, de lo anterior se sigue que un rasgo incrementa la aptitud sólo si hay variación o “competidores” actuales. Sin embargo, los casos de selección y, por lo tanto, de variación actual son escasos. En este sentido, frecuentemente no podemos considerar la función de un rasgo como aquello por lo que es *actualmente* seleccionado contra variaciones *actuales*. Así, en el caso de que el rasgo no sea actualmente seleccionado bajo presiones selectivas de variaciones específicas del rasgo, tenemos entonces que considerar su función en relación al efecto por el que *fue* seleccionado en el *pasado* más reciente.

En cuanto a la crítica de trivialidad explicativa, Millikan (1989) responde que dicha crítica depende de una confusión. Decir que un rasgo R tiene la función F quiere decir que R tuvo una historia durante la cual fue seleccionado por hacer F . Si lo que queremos es saber por qué los miembros actuales de una especie tienen el rasgo R , la respuesta es que R tiene la función F y R fue seleccionado por hacer F . Esta explicación sería vacua si F hubiera sido la función de los primeros ejemplares seleccionados de R . Sin embargo, los primeros ejemplares de R , aunque tenían el efecto que ahora llamamos F , no tenían estrictamente F como función. Podemos llamar E al efecto adaptativo de R antes de que fuera seleccionado, esto es, antes de que se fijara como F . Puesto que el hacer E de R incrementaba la aptitud del organismo más que otras variaciones de R , R fue seleccionado. A partir de entonces E es F . Una función es entonces un efecto adaptativo una vez que es seleccionado. Una vez que esto ocurre R se vuelve una adaptación.

causadas por conductas monógamas pasadas dan cuenta de la existencia de casos actuales de conductas monógamas (p. 117).

⁴⁸ A lo que los valores de aptitud de varios individuos que tienen el rasgo hubiera sido si no lo hubieran tenido, y a los valores de aptitud de varios individuos que no lo tienen hubieran sido si lo hubieran tenido.

Aunque *E* y *F* son el mismo efecto, *E* se vuelve *F* sólo retrospectivamente al episodio de selección relevante. El problema con la crítica es que no distingue entre meros efectos adaptativos y efectos seleccionados (ver (1) arriba).⁴⁹ Si las funciones fueran efectos adaptativos tendría lugar la trivialidad de la explicación etiológica. Puesto que no todos los efectos adaptativos son de hecho seleccionados, no hay peligro de circularidad. Podemos decir sin vacuidad que retrospectivamente el efecto adaptativo devino en función si y sólo si podemos determinar que hubo selección de ese rasgo por dicho efecto.^{50- 51}

Ahora, es cierto que la motivación de Bigelow y Pargetter para desmarcar la teoría de las funciones de la selección natural es que, según ellos, una teoría naturalista aceptable de las funciones biológicas no debe de depender de una teoría particular como es la teoría de la evolución por selección natural. La razón para justificar esta motivación –y que constituye una segunda crítica a la teoría etiológica además de la vacuidad explicativa, es que nuestras intuiciones con respecto al concepto de función, y en particular de función biológica, no dependen de la aceptación de una teoría particular. La idea es que intuitivamente atribuimos

⁴⁹ Bigelow y Pargetter no sólo no distinguen entre meros efectos y funciones de un rasgo, sino que tampoco distinguen entre *tener* una función y *ejecutar* dicha función *precisamente porque para su teoría un rasgo posee la función con tener la mera propensión a incrementar la aptitud aun antes de ser seleccionado*. En la medida en que los efectos son posteriores a sus causas, los efectos si se identifican con sucesos o estados de cosas futuros. Así, la ejecución de una función es un efecto y por lo tanto es un suceso futuro. Pero las funciones *qua* funciones, el *tener* un función, no es lo mismo que ejecutarla. *Ergo*, las funciones no son sucesos o efectos futuros sino efectos seleccionados pasados.

⁵⁰ El problema aparente de circularidad es que si la función *F* de un rasgo *R* explica la persistencia de *R* en la población *P*, entonces no podemos explicar, sin circularidad viciosa, que *R* tenga *F* apelando al hecho de que *R* ha persistido en *P*. Pero puesto que la teoría etiológica distingue entre efectos adaptativos (efectos que aumentan la aptitud) y funciones (efectos seleccionados) –el efecto adaptativo *E* de *R* causó la selección de *R* y luego entonces *E* se volvió *F* y consecuentemente *R* se volvió una adaptación, esta circularidad no es un problema para la teoría: *E más la selección natural* explica la persistencia de *R* en *P* donde *F* es una abreviación para “*E más la selección natural*”. Podemos decir que *F* de *R* explica la persistencia de *R* en *P*, y que *E* de *R más la selección de R por E* explica que *R* tenga *F*. El propensionista podría responder que si *F* es una mera abreviación de “*E más la selección natural*” (*E + SN*) entonces $F = E + SN$, lo cual implica que la distinción entre *F* y *E + SN* no hace ninguna diferencia. Sin embargo, esta objeción es espuria puesto que la diferencia proviene de la estructura temporal histórica de la explicación: puesto que primero fue *E*, luego *E + SN* y sólo entonces $E + SN = F$. Así, *R* no tenía *F* desde el principio por lo tanto, explicar su persistencia apelando a *F* no es circular.

⁵¹ Esta crítica de “vacuidad explicativa” de la visión histórica ha sido también criticada por Mitchel (1993) con la objeción de que Bigelow y Pargetter asumen un adaptacionismo demasiado fuerte. Lo que ella afirma es que si decimos que el rasgo *X* persistió porque tuvo una consecuencia responsable de su selección y evolución consecuente, la única manera en que esto sería vacío es si asumimos que el *único* mecanismo que puede explicar la presencia actual (o persistencia) de un rasgo es la selección natural, cuando en realidad hay otras fuerzas en juego que determinan la evolución (como lo es la deriva genética aleatoria). De ahí que el problema con la crítica de Bigelow y Pargetter a la teoría etiológica selectiva de Millikan es que asumen un adaptacionismo demasiado fuerte, en donde la selección natural es la única causa de la evolución, lo que en verdad volvería a las explicaciones etiológicas “Just-so-stories”.

funciones a rasgos biológicos *independientemente* de cualquier información sobre su evolución o desarrollo e incluso *antes* de que la teoría Darwinista viera la luz. Así, aunque dicha teoría es ampliamente aceptada como verdadera por la comunidad científica relevante, su verdad es *contingente*. En este sentido, si la teoría Darwinista resultará falsa, el recuento etiológico de las funciones se desplomaría con ella.

Para ilustrar este choque entre nuestras intuiciones pre-teóricas y la versión etiológica, Bigelow y Pargetter ofrecen el siguiente experimento mental: Imaginemos un mundo posible *M2* idéntico al actual (*M1*) en que todas las leyes que se cumplen en *M1* se cumplen en *M2* y donde todos los hechos particulares que se dan en *M1* se dan también en *M2* pero con la diferencia de que *M2* empezó a existir, por azar –“sin causa alguna” (p. 104), hace cinco minutos. Dada la teoría etiológica, no habría funciones biológicas en *M2*. En particular, en *M2* habría organismos con rasgos, pero puesto que estos organismos y sus rasgos carecen de la historia-selectiva y del desarrollo adecuado, tendríamos que negar que dichos rasgos tengan funciones. Por ejemplo, los riñones de los organismos en *M2* no tendrían la función de limpiar la sangre aunque eso sea precisamente lo que hagan los riñones en *M2*. Puesto que nuestras intuiciones [ojo, las intuiciones populares o pre-teóricas no las intuiciones de los biólogos] nos inclinan a atribuir funciones a los rasgos de organismos en *M2* a pesar de que *M2* no posee una genuina historia evolutiva, el recuento etiológico resulta radicalmente contra-intuitivo. Debemos concluir que el recuento de las atribuciones funcionales no puede depender de la teoría Darwinista como propone la teoría etiológica.

Para resistir las conclusiones del experimento mental debemos rechazar tanto la pertinencia de dicho experimento como su motivación, a saber, que la teoría naturalista de las funciones biológicas debe ser “independiente” de una teoría biológica particular contingente (ver nota 30). Ofreceré dos razones contra la pertinencia: (1) Dada la teoría biológica evolutiva y dada las intuiciones de los biólogos derivadas de esta teoría, *M2* es biológicamente imposible. (2) El experimento mental no es adecuado en cuanto que *M1* y *M2* no son lo suficientemente similares para constituir un una comparación relevante. La raíz de la motivación es la idea equivocada de que hay una convergencia entre las intuiciones pre-teóricas y las derivadas de aceptar una teoría.

(1) En general, como muestra la historia de la ciencia, nuestras intuiciones ordinarias pre-teóricas sobre el mundo suelen ser radicalmente modificadas o de plano abandonadas una vez que se ha producido una teoría científica al respecto. Las teorías científicas, y por lo tanto, los análisis naturalistas de ciertos conceptos continuos con ellas, no están entonces obligados a capturar en la teoría las intuiciones pre-teóricas relativas a un concepto.⁵² Bigelow y Pargetter se equivocan al incluir en nuestras intuiciones pre-teóricas las intuiciones biológicas y al esperar que las primeras constituyan un argumento contra las segundas. Ciertamente tenemos intuiciones pre-teóricas sobre el mundo orgánico y estas intuiciones motivan el uso del lenguaje funcional en ese contexto e incluso dichas intuiciones pueden jugar un papel en la génesis de la teoría, pero estas no son intuiciones biológicas, científicas. Las intuiciones científicas se derivan de aceptar tal o cual teoría.⁵³ Así, un biólogo evolucionista no tendría que exhibir las intuiciones que Bigelow y Pargetter recomiendan con respecto a su experimento mental. Más bien, un biólogo evolucionista *qua* biólogo evolucionista -no *qua* persona con sentido común, resistiría atribuirle funciones a los rasgos de organismos en *M2* como recomienda la teoría etiológica y la teoría Darwinista. La razón es que la filogenia y la ontogenia de los organismos en *M2* es tan radicalmente anormal que podemos esperar que la respuesta del biólogo evolucionista sería que *M2* es simplemente biológicamente imposible (ver Millikan, 1996) aun cuando pre-teóricamente sintamos la inclinación a atribuir funciones en *M2*.⁵⁴ Así, *M2* es irrelevante para determinar qué

⁵² Por ejemplo, la gente tiene intuiciones sobre lo que es el calor y las ha tenido antes de que cualquier teoría científica al respecto viera la luz. Sin embargo, nadie espera que estas intuiciones sean incorporadas a la teoría cinética-molecular del calor o a cualquier otra teoría sobre el calor. Y nadie se queja de que esto sea un problema para la teoría cinética o para las intuiciones pre-teóricas sólo porque antes de la existencia de dicha teoría la gente tenía intuiciones sobre lo que es calor y dichas intuiciones no son recogidas en la construcción teórica. Más aun, las intuiciones de los científicos que trabajan bajo el amparo de la teoría cinética tienen intuiciones relativas a esa teoría que naturalmente no convergen con las intuiciones pre-teóricas.

⁵³ También es cierto que de hecho la teoría Darwinista recupera las intuiciones pre-teóricas con respecto al “diseño” aparente de los organismos y sus rasgos, esto es, no elimina el componente teleológico sino que da cuenta de él sin recurrir a un diseñador (ver nota 17).

⁵⁴ Millikan (1996) coincide en que la situación planteada en este tipo de experimentos mentales es tan radicalmente improbable que la hace imposible. Si bien ellas están pensando en el experimento del “hombre del pantano” la idea es la misma que el experimento que Bigelow y Pargetter proponen: suprimir la historia. La imposibilidad radica en que no se puede en principio construir moléculas orgánicas grandes como la hemoglobina sólo arrojando las moléculas de amino ácidos de la forma correcta con el ángulo adecuado y con la energía necesaria y hacer que se “peguen”. Dichas moléculas deben ser construidas mediante largas cadenas de reacciones enzimáticas en cascadas que procedan en el orden adecuado y que requieren el soporte del tipo adecuado de catalizadores programados para surgir a escena en el momento preciso. Así, la relevancia de la estructura diacrónica de estos procesos es incompatible con los cinco minutos que lleva de existir *M2*.

puede contar como una atribución funcional legítima *en biología*. En cambio, el experimento sirve para hacer patente la diferencia entre intuiciones pre-teóricas e intuiciones científicas.

(2) Para dar lugar a organismos *biológicos*, no basta que todas las leyes y todos los hechos particulares se obtengan en *M1* y *M2*. La biología, en particular la biología de la evolución, no es una ciencia paradigmáticamente nomológica sino constitutivamente histórica, y la historia esta precisamente excluida en *M2*. Si bien el experimento mental esta diseñado para oponer las intuiciones contra la noción histórica de función, podemos ejemplificar el problema fundamental con esta a-historicidad en torno a la noción de especie. Dada la noción biológica de especie, los organismos en *M2* no pertenecen a una especie –de la misma manera que los rasgos de dichos organismos en *M2* carecen de funciones, pues el principio según el cual miembros de una especie son miembros de la misma clase es radicalmente diferente del principio según el cual un tigre en *M2* y en *M1* pertenece a la misma especie. Ghiselin (1974) y Hull (1978) señalan que el uso biológico de “especie” indica que no se trata de clases espacio-temporalmente irrestrictas sino de entidades históricas que duran en el tiempo; las especies son individuos y los organismos individuales ligados genealógicamente son partes de dichos macro-individuos. A que especie pertenece un organismo individual depende no de sus propiedades a-históricas sino de sus relaciones genealógicas con otros individuos donde dichas relaciones están históricamente circunscritas: Los tigres deben nacer de otros tigres para ser tigres no sólo ser *como* otros tigres.⁵⁵

⁵⁵ Puesto que las especies son entidades históricas (individuos), según Hull (1978) no pueden funcionar en leyes científicas: los individuos no son ejemplos o instancias sino partes de las especies y las partes no tienen que ser similares para ser partes del mismo todo. Así, ningún enunciado de la forma “la especie X tiene la propiedad Y” es una ley de la naturaleza, esto es, una ley universal, verdadera por necesidad natural en lugar de por contingencia histórica. Incluso tampoco es una ley en el sentido *ceteris paribus*: en el caso de las especies la semejanza entre los miembros de una especie, aunque sea estadísticamente significativa, este resultado probabilística no define la membresía de un individuo a una especie sino la genealogía. Así, si entendemos las “clases naturales” en el sentido de Putnam (1975) como definidas en referencia a la naturaleza interna o a las propiedades intrínsecas de una entidad (su microestructura por ejemplo), las especies no son clases naturales. Puesto que algo cae dentro de una clase natural en virtud de su estructura interna, las clases naturales son clases sobre las que se pueden derivar leyes estrictas. Pero la identidad-tipo de tigre o hombre no depende de la estructura interna de estos organismos si no los lazos genealógicos entre individuos *en el tiempo* (lo que Millikan (1984) llama “familias reproductivas establecidas”). Ahora, el tipo de entidades históricas que conforman las especies, si bien no permite derivar generalizaciones nomológicas (ni estrictas ni *ceteris paribus*), si permite derivar generalizaciones históricas como las que operan en la biología de la evolución. Y si bien es posible pertenecer a más de un tipo de clase al mismo tiempo, la idea es que la clase a cuya pertenencia resulta constitutiva de aquello que cae bajo ella es histórica en este caso. Así, hay interacciones lógicas, nomológicas e históricas entre individuos y su pertenencia a ciertas clases donde en cada caso las condiciones iniciales son radicalmente diferentes, en particular, las condiciones iniciales de las generalizaciones históricas no sólo son contingentes sino en muchos casos únicas e irrepetibles. Pero en el caso de

Así, aunque nuestras intuiciones pre-teóricas no inclinan a incluir los tigres en *M2* dentro de la misma especie que los tigres en *M1* por su parecido, las intuiciones biológicas van en otra dirección: no es posible hablar de especies y por lo tanto de membresía a una especie en un mundo que lleva cinco minutos de existencia. Podemos concluir, como en (1), que en ausencia de toda ontogenia y la filogenia no podemos darle sentido a la noción de función o de especie:⁵⁶ lo que se cumple para las especies se cumple para las funciones.^{57- 58}

Finalmente, la teoría propensionalista identifica a las propiedades funcionales con propiedades disposicionales bajo la motivación de que las propiedades históricas no tienen la relevancia causal necesaria en una explicación científica. Pero podemos rechazar que las propiedades históricas que generan el tipo de explicaciones relevantes en la biología de la evolución no sean causalmente eficaces en comparación con las explicaciones disposicionales, y por tanto rechazar la identificación de funciones y disposiciones. ¿Son los efectos seleccionados pasados de ciertos rasgos causalmente inertes con respecto al presente? Ciertamente no son propiedades causales actuales (sincrónicas) de los rasgos, pero si son propiedades causalmente

las especies sólo las interacciones históricas son constitutivas de lo que es pertenecer a una especie. Lo mismo para las funciones, sólo que, si bien la genealogía es relevante, el elemento central es la selección natural.

⁵⁶ Si bien puede haber mundos posibles donde existen “especies” pero donde no haya historia, en dichos mundos la noción de “especie” es radicalmente diferente a la noción biológica de especie en el *M1* como para que podamos concluir que el término “especie” tiene significados radicalmente diferentes. Quizás los mundos posibles donde hay especies sin historia son también mundos donde hay una ciencia que se llama “biología evolucionista” y que se avoca a dar cuenta de qué son las especies. Sin embargo, dicha ciencia es también radicalmente diferente de lo que llamamos biología en *M1*.

⁵⁷ Millikan concluye algo similar con respecto al “hombre del pantano”.

⁵⁸ En cuanto al temor de fundar el recuento de las funciones biológicas en una teoría científica contingente particular, podemos preguntar, si lo que queremos es hacer compatibles las explicaciones que apelan a funciones con explicaciones científicas naturales, ¿cómo hacer esto sin apelar a una teoría o teorías científicas particulares? Más aun, ¿qué teorías particulares en ciencia empírica no son contingentes? Tal parece que el precio por ofrecer una teoría naturalista, en este caso de las funciones biológicas, es comprometerse con alguna o algunas teorías científicas particulares o al menos con ciertos elementos de ellas. Podemos disolver esta crítica a la propuesta etiológica de la siguiente manera. Primero, al menos fuera de las matemáticas y la lógica, todas las teorías científicas son, si verdaderas, contingentes y como tal pueden ser falsas (quizá *deben* poder ser falsas *sensu* Popper (1951). Dada la metodología empírica, una teoría *necesariamente* verdadera perdería, *ipso facto*, su estatus científico-naturalista. Segundo, las teorías científicas no son falsas o verdaderas *simpliciter*, sino que están sujetas a modificaciones en las que ciertos elementos de la teoría se conservan mientras otras simplemente se abandonan. La historia de la ciencia está llena de ejemplos: la mecánica newtoniana con respecto a la mecánica cuántica, la teoría de la evolución de Darwin, la teoría de Mendel, y en particular la “síntesis” de estas dos. Tercero, el abandono de una teoría particular *T1* sólo tiene sentido en relación a *otras* teorías alternativas, *T2, T3, ... Tn*, que idealmente no sólo expliquen todo lo que *T1* explica, sino que den cuenta de las anomalías en *T1* y que sean capaces de explicar nuevos hechos que no están incluidos en *T1* (Lakatos*) ¿Qué alternativa a la teoría Darwinista mencionan Bigelow y Pargetter? ¡Creacionismo! Sobra decir, (aun así lo diré) que esa *no* es una teoría *científica*, por lo tanto, no es una alternativa genuina o relevante a la teoría Darwinista, teoría que es ampliamente reconocida por haberle dado el tiro de gracia a esa “explicación” teológica.

eficaces consideradas diacrónicamente (ver nota 47): casos pasados de efectos de un rasgo (en parte) causan casos presentes de dicho rasgo y por lo tanto (en parte) causan casos presentes de dichos efectos. Si bien la causa *actual* de la presencia de un rasgo caso es la ontogenia del organismo portador, podríamos decir que la ontogenia recapitula la filogenia, de ahí el “en parte”.⁵⁹ - ⁶⁰ Si bien es cierto que tanto las propiedades funcionales como las propiedades disposicionales son múltiplemente realizadas, la diferencia estriba en que las disposiciones son propiedades *actuales* de las estructuras sobre las que sobrevienen (sus bases categoriales) mientras que las funciones son propiedades *pasadas* (históricas) de las estructuras sobre las que sobrevienen. Y esta diferencia es constitutiva: las propiedades funcionales sobrevienen sobre las estructuras morfológicas consideradas históricamente. Es en este sentido que distintas estructuras biológicas pueden tener la misma función y, sin embargo, tener diferentes disposiciones. Por ejemplo, las alas de los murciélagos y de las aves tienen la función de volar, pero realizan esto mediante distintas disposiciones como es evidente dada la diferencia entre las estructuras causales de las alas de una especie y la otra. Esto implica que un rasgo puede *tener* una función y, sin embargo, carecer de las disposiciones relevantes para *ejecutarla*. Podemos decir que una función ejecutada normalmente es idéntica a las disposiciones correspondientes de la estructura biológica en cuestión, sin embargo, tener una función y ejecutarla son propiedades diferentes, si bien en el origen selectivo coincidieron. Además, esta diferencia entre *tener* una función y *ejecutarla* es la base para hablar de la normatividad de las funciones.⁶¹ Si identificamos a las funciones con disposiciones, ¿cómo podemos explicar los casos de disfuncionalidad? Al parecer no podemos, a menos de que recurramos a alguna suerte de consideración histórica. Así, las funciones no son disposiciones, pues las funciones son propiedades históricas.

⁵⁹ De la misma manera que un caso anterior de un efecto *E* es capaz de afectar el rol causal que casos posteriores de *E* jugaran en la vida de un organismo (ontogénesis), así también casos ancestrales de *E* afectan el rol causal que casos futuros de *E* jugaran en generaciones subsecuentes (filogénesis).

⁶⁰ Además, las explicaciones evolutivas son históricas, y como tales no sólo tiene poder predictivo sino poder retrodictivo. El que dicha teoría realice retrodicciones no se explica como una deficiencia de la teoría sino como una condición intrínseca al fenómeno que estudia, a saber, la vida y su evolución. De la misma manera, nadie negaría el estatus científico de la astronomía sólo porque genera explicaciones sobre sucesos pasados. Decir que lo que hace a una explicación científica ser “genuinamente” científica (Bigelow y Pargetter) es la capacidad de derivar predicciones, es sin duda un sesgo positivista que eclipsa otras virtudes epistémicas de las teorías científicas.

⁶¹ Ésta es la asimetría entre lo que se supone que un rasgo debería hacer, dada su historia evolutiva, y lo que de hecho hace dadas las condiciones actuales de ejecución.

5. Teoría de la historia moderna.

Godfrey-Smith (1994) propone una versión modificada de la teoría etiológica en la que el episodio de selección relevante para la individuación de una función es el históricamente más reciente. Así, las funciones son efectos que explican el *mantenimiento reciente* de un rasgo en una población. Esto quiere decir que no cualquier efecto pasado constituye una explicación de la función actual. La razón es que la selección natural puede intervenir en la *producción* de un rasgo, pero también en el *mantenimiento* posterior de dicho rasgo en una población frente a posibles variaciones que surgen de la recombinación del material genético.⁶² Así, el episodio de selección original puede estar dirigido a un efecto distinto al episodio de selección más reciente y por tanto el mismo rasgo puede haber tenido distintas funciones a lo largo de su historia evolutiva.⁶³ Esta distinción entre “historia reciente” e “historia antigua” pretende reflejar la distinción entre una explicación funcional y una explicación evolutiva tal y como ocurre en la práctica biológica. Mientras la exigencia de explicación evolutiva responde a la pregunta por las fuerzas que originalmente dieron lugar a un rasgo, la demanda de explicación funcional responde en cambio a la pregunta por el mantenimiento reciente del rasgo.⁶⁴ Una teoría etiológica adecuada, según Godfrey-Smith, debe hacer patente esta distinción.

La teoría de la historia moderna, si bien reconoce el elemento histórico y selectivo para la determinación de las atribuciones funcionales, parte de reconocer que hay un elemento *no* histórico relevante en las atribuciones funcionales que los biólogos realizan, a saber, que en la práctica biológica los biólogos suelen reservar el término “función” para actividades o fuerzas

⁶² Cabe preguntar al modernista, ¿Qué tan *reciente* debe ser el pasado para contar como un episodio selectivo *relevante*? Para Godfrey-Smith esto no constituye un problema serio: “*relevance fades*” (1994, p. 214). La razón que da es que los episodios de selección se vuelven cada vez más irrelevantes para la adscripción de funciones a un tiempo específico en la medida en que nos vamos más lejos en el tiempo.

⁶³ La distinción que la teoría moderna introduce está motivada por la distinción entre “adaptación” y “exaptación” de Gould y Vrba (1982). La idea es que, mientras las adaptaciones son caracteres moldeados por la selección natural para el rol que ejecutan en la actualidad, las exaptaciones son caracteres construidos originalmente por la selección para un rol o caracteres con ninguna explicación selectiva del todo, que desde entonces han sido cooptadas para un nuevo uso. En este sentido sólo las adaptaciones tienen funciones mientras que las exaptaciones tienen sólo efectos. La idea central es dar cuenta de la importancia de distinguir entre el uso original de un rasgo y su uso corriente.

⁶⁴ Es importante recordar que puesto que la selección natural es *una* de las fuerzas operando en la evolución junto con otros procesos estocásticos (como la deriva genética o la mutación), esta explicación evolutiva del origen no necesariamente tiene que ser una explicación *selectiva*. De hecho, tal explicación evolutiva puede ser diferente de la explicación de por qué tal rasgo ha sido recientemente mantenido dentro de la alberca genética de una población.

que, en algún sentido intuitivo, son útiles o constructivas (Godfrey-Smith, 1994). La consecuencia es entonces que sería deseable hacer alguna restricción en los procesos selectivos que son relevantes para determinar el estatus funcional de un rasgo, de tal suerte que no todo incremento en la aptitud cuente como función.⁶⁵

Podemos resumir la reforma que la teoría de la historia moderna hace a la formulación histórica original de Millikan de la siguiente manera: (1) la contribución general que un tipo de rasgo ha tenido con respecto a la aptitud en el *pasado* no determina su contribución *actual* a la aptitud; (2) la contribución actual a la aptitud debe ser al menos parte del concepto de función evolutiva; y (3) el problema con la teoría etiológica es que este elemento no es histórico, algo pasado, sino actual. No es lo mismo explicar la *producción* de un rasgo que su *mantenimiento* en una población. Esto equivale a confundir la utilidad *original* (histórica) de un rasgo, con su utilidad *actual*. La función por la que un rasgo evolucionó en el pasado no necesariamente es la función actual del rasgo. En este sentido, la explicación etiológica parece identificar la explicación *evolutiva*, que es histórica, con la explicación *funcional*, que no lo es necesariamente.⁶⁶

El problema con la versión original de Millikan es que presenta a la explicación etiológica como una explicación histórica, excluyendo así otras fuentes de asignación funcional en biología. La idea es entonces que las explicaciones selectivas empleadas en las explicaciones

⁶⁵ Cabe señalar que otras críticas similares se han hecho a la formulación de Millikan. Para Walsh y Ariew (1996), la teoría histórica o etiológica es incompleta y comete el error de pensar a todas las funciones biológicas como funciones históricas. La idea central contra la restricción histórica de la teoría etiológica es que la utilidad *pasada* no determina la utilidad *actual*. Walsh y Ariew ponen el ejemplo de la fisiología y la anatomía funcional como casos en los que si un tipo de rasgo en general contribuye a la aptitud de forma novedosa en una población, tal forma novedosa constituye un caso legítimo de atribución funcional. La teoría etiológica se limita a la utilidad pasada y, por tanto, no puede dar cuenta de las funciones *nuevas* de los rasgos que incrementan la aptitud de los portadores en el presente.

⁶⁶ Hay una diferencia importante entre los motivos detrás de la teoría de Millikan y la de Godfrey-Smith. Para ella el objetivo es mostrar los elementos en común entre adscripciones de funciones propias en contextos biológicos, artefactuales y psicológicos y en específico ofrecer una descripción naturalista de las atribuciones de propósito y de la intencionalidad en particular. De hecho, Millikan (1984, 1993) explícitamente afirma que no pretende hacer un análisis conceptual del uso de “función” en biología, sino que trata de introducir un término teórico que sirva a sus propósitos, “función propia”. Por supuesto, la idea de Millikan es que si bien no pretende capturar el uso de los biólogos, su teoría de las funciones propias sí pretende iluminar en qué sentido tales atribuciones son teleológicas. Esto es de esperarse ya que si la naturalización de la intencionalidad se realiza apelando a funciones teleológica evolutivas, dicha noción de función tiene que estar sustentada *en* la biología evolucionista. Por su parte, Godfrey-Smith sí pretende capturar el uso biológico de “función” y como tal apela a la práctica biológica para mostrar la relevancia de distinguir entre explicaciones evolutivas o históricas y explicaciones funcionales. Sin embargo, Godfrey-Smith (1996) también extrae consecuencias para entender el lugar de la mentalidad en el mundo natural a partir de su análisis de nociones teleológicas o “teleonómicas”.

etiológicas deben ser las explicaciones selectivas genuinas de los rasgos en cuestión pero, además, que en algunos casos la explicación selectiva y la funcional pueden divergir. No obstante, desde su obra seminal (1984) Millikan explícitamente reconoce explicaciones normales más o menos próximas en el tiempo además de que explícitamente admite otro tipo de funciones en biología (1993). Desde 1984, ella incluye en su explicación la distinción entre función y *vestigio* en contraste con otras versiones como la de Neander (1991). La idea de Millikan (1984, p. 32) es que el patrón histórico de correlación que es responsable de la existencia de una entidad o estructura que posee una función propia puede ser más o menos próximo al tiempo de la producción de dicha entidad. En este sentido, una función propia es “históricamente próxima” o “históricamente remota”. El ejemplo de un vestigio es el apéndice humano, cuya correlación histórica entre función y existencia resulta remota en tiempo. Esto explica que la función propia de nuestros apéndices sea históricamente remota. Si la función que da cuenta de la reproducción de sus ancestros a través del tiempo cambia, la entidad o estructura acumula una serie de funciones propias donde cada una es más próxima a los casos actuales que la última. Millikan además menciona que en la biología no abundan estos casos en virtud de que *la estructura de un órgano suele cambiar cuando su función cambió*, por ejemplo, las piernas delanteras se convierten en aletas.⁶⁷

⁶⁷ Griffiths (1993) desarrollar la distinción entre rasgos funcionales actuales y rasgos vestigiales. La idea de vestigialidad está ligada a la noción de evolución regresiva, esto es, el proceso mediante el cual rasgos inútiles tienden a atrofiarse debido a los costos de producirlos. Pero no todos los casos de vestigialidad corresponden a la evolución regresiva. Griffiths (1993) señala dos. Primero, un rasgo puede cambiar su función pero ser preservado en virtud de que ejecuta una nueva función. Segundo, un rasgo inútil puede ser preservado por un largo periodo de tiempo simplemente debido a que a no hay variación genética. En este sentido, no todos los rasgos que cesan de ejecutar su función deben ser clasificados como vestigios ya que un rasgo disfuncional no puede ejecutar su función propia y sin embargo, la tiene. La propuesta de Griffiths (1993) para incluir la vestigialidad en el contexto de una teoría de las funciones propias supone; a) definir un periodo evolutivo de tiempo significativo para un rasgo T como un periodo tal que, dado un rango de mutación en el *loci* que controla T y el tamaño de la población, podríamos esperar que suficiente variación de T haya ocurrido tal que permitió una evolución regresiva significativa si el rasgo no hacia ninguna contribución a la aptitud; b) un rasgo es un vestigio relativo a alguna función pasada F si no ha contribuido a la aptitud al ejecutar F por un periodo de tiempo evolutivo significativo; c) un rasgo es un vestigio *simpliciter* si es un vestigio relativo a todas sus funciones pasadas. Esta propuesta permite que un rasgo se convierta en un vestigio relativo a una función específica y, sin embargo, mantenerse intacto en virtud de otras funciones suyas. Al mismo tiempo, permite que un rasgo se convierta en un vestigio y, no obstante, mantenerse intacto debido a ausencia de mutación genética y permite que un rasgo sea un vestigio para una población sin tener que serlo para otra. Finalmente, para Griffith (1993) este uso concuerda con el uso biológico.

6. Conclusión: Funciones de rol causal y el pluralismo des-unificador.

La clase de funciones que he revisado corresponde las funciones evolutivas, pero éstas no son la única clase de funciones empleadas en las ciencias biológicas. Introduciré otra noción de función y expondré qué relaciones hay entre ambas nociones. El propósito de este apartado es hacer patente la relevancia de las funciones evolutivas en el contexto más amplio de este ensayo, a saber, el problema de la naturalización de la intencionalidad así como defender una posición pluralista con respecto al rol de las atribuciones funcionales en la biología.

Las funciones no teleológicas o de rol causal son asignadas en un análisis de una capacidad compleja de un sistema, al descomponer dicha capacidad en un conjunto de capacidades más simples. Estas capacidades simples, a su vez, pueden ser analizadas en términos de capacidades más simples, hasta el nivel en el que las capacidades pueden ser directamente explicadas por las leyes de la física (Cummins, 1975, 1983, 2002). Diferentes descomposiciones se pueden hacer de un mismo sistema dependiendo de consideraciones pragmáticas. La función de una entidad es su rol causal, su contribución disposicional a la capacidad general de un sistema contenedor. Se trata de explicaciones mecánicas abstractas (Cummins, 2002) en las que la entidad es una parte funcional del sistema. La capacidad general se explica mediante una descomposición de las capacidades que contribuyen como partes de un sistema contenedor.

Definición (df): dado un sistema s (organizado) funcionalmente, la función de x en s es la de hacer G relativamente a un análisis de A de la capacidad de s de hacer y sólo en el caso en el que x es capaz de hacer G en s y A da cuenta adecuada y apropiadamente de la capacidad de s de hacer y mediante, en parte, apelar a la capacidad de x de hacer G en s (Cummins, 1975, p. 76).

Las funciones causales se asignan al tratar de explicar *cómo* funciona alguna capacidad de un sistema, esto es, cómo las partes del sistema interactúan funcionalmente entre sí. La explicación apela a propiedades que el sistema tiene en virtud de sus disposiciones causales actuales, independientemente de la historia del sistema y de los detalles del medio. La individuación de estas funciones es internista (Amundson y Lauder, 1993, Cummins, 2002).

Ahora bien, en principio estos tipos de funciones biológicas difieren en dos sentidos. Primero, en el tipo de explicaciones que genera cada tipo de atribución funcional, la explicación

etiológica y el análisis funcional. Segundo, en las propiedades funcionales que son asignadas por cada tipo de atribución. En el caso etiológico la propiedad relevante la constituyen efectos seleccionados pasados, mientras que en el otro, son efectos actuales independientes de todo proceso de selección, *ergo*, independientemente de toda referencia al contexto histórico y ecológico. Aunque este pluralismo es en general reconocido en la práctica, los filósofos de la biología han propuesto posibles relaciones entre ambos tipos de función. Podemos destacar tres:

- (i) Un tipo de función *depende* de otro por lo que pueden ser *unificados*, ya sea subsumiendo el modelo etiológico en términos de un análisis funcional previo de la capacidad en cuestión (Griffiths, 1993) o explicando las atribuciones de rol causal en términos de una “fuente de diseño” a la base de la atribución (Kitcher, 1993).
- (j) Los dos tipos son *independientes* entre sí (Millikan, 1989, Godfrey-Smith, 1993) puesto que responden a diferentes exigencias de explicación legítimas en biología dirigidas a dar cuenta de distintas propiedades funcionales.
- (k) La eliminación de un tipo a favor del otro (Cummins, 2002).

Dejaré de lado (iii) para concentrarme en (i) y (ii). (i) implica que la *validez* de un tipo de atribución funcional depende del otro. La idea es que primero hemos de completar un análisis funcional del rasgo o capacidad en cuestión para poder posteriormente inferir su etiología (Griffiths, 1993). Según la propuesta unificadora de Griffiths, el poder explicativo de las funciones que son efectos seleccionados proviene no directamente de la etiología del rasgo, su historia causal, sino de sus disposiciones actuales que dan cuenta de cómo dicho rasgo ejecuta la capacidad correspondiente. Lo que afirma es que un análisis funcional es necesario para la etiología mientras que el análisis funcional es independiente de la etiología. Me parece que esta dependencia asimétrica debe ser entendida en un sentido más robusto como negando todo genuino poder explicativo a la etiología *per se*. En este sentido, la versión de Griffiths (1993) parece más cerca de (iii) que de (i), al menos si pensamos la legitimidad de las propiedades funcionales en términos de la legitimidad de las explicaciones en donde las atribuciones correspondientes ocurren. Lo que lleva a esta consecuencia más radical es que, dada la

dependencia asimétrica, las funciones que son efectos seleccionados no son explicativas consideradas independientemente de otras explicaciones o de otros tipos de función.

Me parece que lo que motiva esta propuesta unificadora es una interpretación mecanicista de lo que una explicación científica o natural debería de ser. La pura etiología no es genuinamente explicativa, porque las explicaciones históricas (y funcionales) deben ser sustituidas, reducidas o analizadas en estrictos términos mecánicos (Stotz y Griffiths, 2002), esto es, en términos que excluyen tanto la historia como la utilidad.⁶⁸ Sin embargo, podemos apelar a razones pragmáticas para concluir que ambos tipos de explicación pueden ser llevados a cabo, en principio, uno independientemente del otro, sin que esto implique que sean excluyentes. La idea central es que responden a dos demandas diferentes de explicación que convergen en el largo plazo. En particular, una explicación normal completa de un efecto seleccionado que incluya los detalles del mecanismo o estructura funcionalmente caracterizada, converge con explicaciones disposicionales y legaliformes como la que Cummins propone (Millikan, 1993). Sin embargo, tanto el carácter relacional como el carácter normativo que interesa destacar de las funciones teleológicas no están presentes en el tipo de atribución funcional que Cummins propone. El “análisis funcional” de Cummins, hace de la función de una entidad su mera contribución causal a un proceso complejo, o a la capacidad general del sistema contenedor. Aunque ciertamente esta versión logra evitar apelar a algo más que la causación eficiente, parece hacerlo con el costo de negar todo contenido teleológico al concepto de función. Esto no sería un problema sino fuera porque esto descarta muchos de los empleos del concepto de función en la práctica y en el discurso biológico. Bajo el análisis de Cummins, no queda descartado que la función del corazón sea la de hacer cierto ruido, por ejemplo. En cambio, las funciones etiológicas son *relacionales* en tanto que la atribución de una función a un rasgo no puede realizarse sin mencionar los detalles del medio ambiente y la historia particular en la que la función ha sido ejecutada. Esto implica que las funciones teleológicas son individuadas de manera *externista*. Finalmente, estas funciones son, como hemos visto, *normativas*, esto es, el atribuir una función teleológica a una entidad permite discriminar ejecuciones correctas de incorrectas de dicha función, haciendo posible la disfuncionalidad.

⁶⁸ Los biólogos realizan de hecho explicaciones mediante atribuciones de funciones selectivas a rasgo, independientemente de un análisis causal del rasgo en cuestión y de la capacidad en cuestión.

CAPÍTULO SEGUNDO: Representación Mental y Teleonomía

“Para decir qué es un significado, podemos primero preguntarnos qué hace un significado y luego encontrar algo que lo haga.”

David K. Lewis, 1970.

1. Introducción.

Una ráfaga de viento sobre la arena puede, por azar, producir una configuración en la arena que podría ser confundida con una huella. Pero *no* es una huella. La *forma* de una marca en la arena no basta para que sea una huella. Lo que hace ser una huella a una configuración de arena es la relación entre esa configuración o forma y el proceso que la produjo en ese contexto, su origen. Muchos estados o propiedades de ciertos objetos son resultado de la relación entre el objeto y su entorno, incluidas las relaciones pasadas con su entorno. Este tipo de estados o propiedades son “amplios” o “gruesos”, sensibles al contexto. La teoría que revisaré afirma que el contenido de una representación mental, como una huella en la arena, es resultado de su historia causal. La propuesta teleosemántica de Millikan (1984, 1993, 2004) es que los estados mentales intencionales –actitudes proposicionales, percepciones, son este tipo de estados y que las propiedades semánticas de estos estados son relacionales en el sentido histórico-causal. El contenido mental no sobreviene en el organismo, o en la cabeza del organismos considerado al margen de su situación ecológica pasada y presente. La *forma* de una representación no basta para explicar su contenido, es necesario dar cuenta de su *función*. Llamaré a esto externismo ecológico.

En la introducción a esta tesis caractericé la intencionalidad y su relación con el problema del estatus naturalista de las propiedades semánticas de las representaciones mentales. Retomaré algunos puntos centrales para motivar la respuesta teleonómica o etiológica a éste problema. En el segundo apartado presentaré el argumento general de la teleosemántica y las intuiciones que la motivan. En el tercer apartado expondré cómo la teoría etiológica de las funciones propias de Millikan (1984) se extiende al campo de la cognición y, en concreto, a la psicología intencional

(creencias y deseos). En el cuarto apartado comenzaré motivando el problema normativo y de la disyunción para las teorías naturalistas de la representación mental para después presentar la respuesta teleonómica a dicho problema. Finalmente, concluiré con un análisis de algunas consecuencias metafísicas del naturalismo Darwinista sobre el lugar de la intencionalidad y lo mental en el mundo natural. A continuación presentaré brevemente y, en líneas generales, qué son las ‘actitudes proposicionales’, la idea de que el contenido de estos estados consiste en sus condiciones de verdad y la relación de éstas con la putativa normatividad de la intencionalidad. También diré algo sobre el sentido en que estas condiciones individúan a estos estados y el papel que juegan en las explicaciones psicológicas. Finalmente, diré algo sobre en qué sentido cabe hablar de estados intencionales en organismos no humanos. En cada caso motivaré el problema naturalista.

¿Qué es la intencionalidad?, ¿qué clase de propiedad es la de tener contenido o significado?, ¿cómo un estado interno de un organismo puede representar o ser un símbolo de otra cosa? La intencionalidad hace referencia a la clase de estados mentales que poseen contenido representacional, las “representaciones mentales”.⁶⁹ Dentro de esta clase de estados mentales podemos incluir a las percepciones y las “actitudes proposicionales” como son las creencias y los deseos. Me concentraré por ahora en estas últimas.⁷⁰ El contenido de una creencia es cómo representa al mundo tal que la creencia es individuada a partir de lo *que* representa y de *cómo* lo representa. Paralelamente, el contenido del deseo ‘que Q’ es el objeto proposicional, el significado de ‘que Q’, que expresa cómo quisiera –desde la perspectiva del agente– que fuera el mundo. ¿Cómo entender esta capacidad representacional del contenido de estas actitudes?

El contenido de una representación mental es la propiedad que determina qué condiciones en el mundo son relevantes en la determinación del valor de verdad de la representación. En otras palabras, el contenido de una representación mental es la propiedad responsable de que el estado

⁶⁹ En general, ‘estados intencionales’, ‘estados cognitivos’, ‘estados representacionales’ o ‘representaciones mentales’ las emplearé como sinónimos.

⁷⁰ Como su nombre lo indica, se trata de actitudes que los agentes toman hacia proposiciones y que pueden ser descritas usando oraciones indicativas en cláusulas oracionales del tipo ‘que P’, e.i. x cree que P, x desea que Q. Por ejemplo, si tomamos la actitud de aceptación con respecto a la proposición de que hace mucho frío, entonces tenemos la creencia de que hace mucho frío. Si tomamos la actitud de desear que la proposición de que hace mucho frío sea verdadera, entonces nuestro deseo es que haga mucho frío. Así, el contenido u “objeto intencional” de una actitud proposicional es un “objeto proposicional”, el significado de la cláusula oracional ‘que P’ en enunciados de atribución del tipo ‘x cree que P’.

pueda ser verdadero o falso en el caso de las creencias, o satisfecho o insatisfecho en el caso de los deseos. Puesto que la propiedad relevante determina qué condiciones en el mundo hacen a la creencia verdadera o falsa o en las que un deseo es satisfecho o insatisfecho decimos que esta propiedad es *semántica*. Así, tener contenido es una condición necesaria de posibilidad para tener un valor de verdad, mientras que la condición suficiente proviene de cómo sea la realidad.⁷¹

Un elemento central es que una representación mental retiene su contenido independientemente de que dichas condiciones se den o de que podamos determinar que dichas condiciones de hecho se dan. O lo que es lo mismo, las representaciones falsas y vacías también tienen contenido. En este sentido, un problema para el naturalista es decir qué tipo de propiedad natural podría efectivamente mediar entre el vehículo de la representación mental y lo representado (en el mundo), tal que la mediación se obtenga independientemente de si de hecho la representación (vehículo) y lo representado (mundo) coinciden; esto es, independientemente de si la representación es verdadera. La idea es que cuando una representación mental es verdadera, en algún sentido se da una relación real entre la representación y lo representado, mientras que cuando una representación mental es falsa, parecería que la relación se da entre la representación y algo inexistente o no actual, puesto que aun así retiene su contenido

Hablar de los pares “verdadero / falso”, “satisfecho / insatisfecho” es hablar de casos específicos de la distinción “correcto / incorrecto”. En este sentido, las propiedades semánticas de las representaciones mentales son propiedades *normativas*; el contenido de una representación expresa las *condiciones de corrección* de la representación. La idea es que es constitutivo del contenido de una representación mental que lo que determina su contenido particular determina cómo emplear correctamente la representación de suerte que podamos *evaluar* el empleo de la representación –su valor de verdad– a partir de su contenido. En el caso de representaciones mentales como las creencias podemos hablar de dichas condiciones en términos de “condiciones de verdad”, mientras que en el caso de los deseos hablamos de “condiciones de satisfacción”. La pregunta es, ¿qué tipo de propiedad natural confiere a estos estados mentales la capacidad de ser

⁷¹ Y, presumiblemente, hay representaciones con contenido que no tienen valor de verdad determinado, ya sea por que son vacías –carecen de referencia–, o porque simplemente no contamos con una manera de determinar si las condiciones se dan o no. Por otra parte, no todos los estados intencionales son acerca de aspectos en la realidad sino que pueden ser acerca de otros estados intencionales, en cuyo caso la condición suficiente no es de la realidad extra-mental.

semánticamente evaluables, esto es, evaluables en relación a su valor de verdad?; ¿qué propiedad natural hace posible que algo este sujeto a condiciones de corrección, ya sean de verdad o de satisfacción?

Si bien la creencia de que el agua del mar es salada y la creencia de que Guatemala está al sur de México tienen el mismo valor de verdad, ambas creencias difieren en las condiciones bajo las cuales son verdaderas y, por tanto, difieren en su contenido. Así, no sólo es necesario dar cuenta de la propiedad que hace posible que algo esté sujeto a condiciones de verdad, esto es, que sea semánticamente evaluable, sino además es necesario explicar cómo los estados mentales *particulares* tienen las condiciones de verdad particulares que tienen. Una teoría del contenido representacional de los estados mentales intencionales debe dar cuenta del modo en que los estados particulares son *individuados* a partir de su contenido. En este sentido, una teoría que explique en términos naturales las “condiciones de corrección” de las representaciones mentales debe poder al mismo tiempo explicar cómo dichas condiciones individúan a las representaciones particulares.⁷²

Si bien las condiciones de corrección o de verdad de las representaciones mentales son constitutivas de la noción de contenido, la idea misma de contenido está directamente vinculada con la explicación de la conducta de los agentes. El contenido de una creencia o deseo es una propiedad causalmente explicativa de los estados mentales en tanto estos estados causan conductas específicas. Decimos, por ejemplo, que el hecho de que Abraxas se haya arrojado por

⁷² Es necesario hacer una pequeña digresión para puntualizar el problema que me ocupa, distinguiéndolo de otros problemas relacionados. Si bien he identificado el contenido de los estados intencionales o de las representaciones mentales con sus condiciones de verdad, la característica opacidad de los contextos de atribución de estos estados muestra que estos estados están más finamente individuados de lo que sus condiciones de verdad hacen patente. La “opacidad” de las actitudes proposicionales refiere al hecho de que la atribución de actitudes proposicionales es sensible a la manera en que los elementos de la atribución son descritos, esto es, a *cómo* son las cosas representadas y no sólo a *qué* representan. Por ejemplo, x puede creer que Ringo Star es el baterista de los Beatles, sin creer que Richard Starkey es el baterista de los Beatles a pesar de que ambos nombres refieren al mismo objeto, en este caso a una persona. Dicho de otra manera, x cree que P y no que Q aunque ‘ P si y sólo si Q ’ sea verdadera. El problema es que si individuamos el contenido de estas creencias en términos de sus condiciones de verdad, y P y Q son verdaderas bajo las mismas condiciones, entonces P y Q tienen el mismo contenido. Sin embargo, como muestra el ejemplo, individuar el objeto proposicional de la actitud en términos de sus condiciones de verdad no permite dar cuenta de esta diferencia en los contenidos de los estados de x . Lo que me interesa aquí es qué tipo de relación hay entre x y P , tal que x se representa P , y tal que la palabra ‘cree’ -que parece denotar algún tipo de relación *sui generis*-, pueda ser reducida a algún tipo de relación natural. Así, nuestro problema no es el de explicar la característica opacidad de los contextos de *atribución* intencional. Ahora, esto no implica que la noción de contenido o condiciones de corrección que estoy empleando sean puramente extensionales. Esto se hará patente en el tercer capítulo.

la ventana se explica por su creencia de que ésa era la única manera de escapar del incendio junto con su deseo de no querer morir quemado. La psicología de creencias y deseos afirma que el comportamiento de los sujetos con creencias y deseos lleva a la satisfacción de sus deseos si sus creencias son verdaderas. Puesto que el contenido determina las condiciones de verdad de un estado intencional, el contenido debe jugar algún papel causal en la generación de acciones (ver la Introducción apartado segundo).

Más allá de las actitudes proposicionales y de la psicología intencional humana podemos hablar *prima facie* de intencionalidad o contenido representacional en contextos no humanos. La mayoría de los organismos poseen sistemas de control con la capacidad de producir estados que podemos describir como estando “dirigidos a” aspectos relevantes de su entorno de tal suerte que puedan producir comportamientos específicos en respuesta a estos aspectos. Por ejemplo, bacterias de cierto tipo en el polo norte poseen pequeños magnetos internos que les permiten alejarse de la superficie del agua rica en oxígeno que les resulta letal; la danza de la abeja representa la localización de un néctar a su audiencia;⁷³ el castor azota su cola para avisar a sus congéneres de la presencia de un depredador; la rana es sensible a objetos pequeños en movimiento en su campo visual que constituyen alimento. Estos son sólo algunos ejemplos de la variedad y la ubicuidad de representaciones en el reino animal. Ahora bien, ¿en qué sentido estas representaciones son intencionales?

En primer lugar estos estados, sean internos o externos, son “acerca de” otra cosa y como tal son representacionales. Más aún, podemos decir cuál es el contenido de una danza particular en el caso de las abejas. Por otra parte, una característica central de las representaciones intencionales es que el error o la falsedad son posibles. En este sentido, los ejemplos arriba señalados son estados capaces de representar equivocadamente, esto es, son estados sujetos a condiciones de corrección y como tal, resultan semánticamente evaluables: casos de las representaciones relevantes pueden ser empleadas bajo otras condiciones produciendo una falta de coincidencia entre el contenido de la representación y su objeto. Si bien en el contexto humano

⁷³ Las abejas danzantes pueden ser también caracterizadas como manteniendo una relación intencional con un elemento ecológicamente relevante en su entorno. La diferencia es que estos insectos sociales, careciendo de un sistema nervioso adecuado, no median estas relaciones intencionales con estados internos representacionales sino que simplemente *reaccionan* adecuadamente al signo externo expresado por la danza. No obstante, la propuesta es que su acción consecuente de dirigirse en la dirección señalada por la danza puede ser descrita intencionalmente, ya que es posible que la representación externa, la danza, indique una localización falsa.

el principal reto para el naturalista es explicar en qué consiste esta normatividad sin apelar a un agente o a alguna noción que presuponga intencionalidad, en el contexto no humano esto es simplemente inevitable. Dar cuenta de esta normatividad es el principal propósito de una teoría teleonómica del contenido intencional que analizaremos puntualmente en el apartado cuatro. Lo importante por ahora es que la noción de intencionalidad o de representación puede ser extendida al contexto animal teniendo en cuenta que debemos reconocer distintas maneras de representar y distintos grados de sofisticación en las capacidades representacionales de los organismos, desde los estados de la bacteria hasta las actitudes de la psicología humana.⁷⁴

Dadas las características arriba señaladas que conforman el fenómeno de la intencionalidad o de la representación mental, podemos circunscribir el proyecto naturalista a las siguientes preguntas o aspectos:

- (a) La pregunta por el *vehículo* de representación: ¿Qué son las representaciones mentales?, ¿mapas internos, símbolos internos, estructuras neuronales distribuidas, oraciones en un “lenguaje del pensamiento”, mapas internos?
- (b) La pregunta por la naturaleza del *contenido* de la representación.⁷⁵ ¿Cómo representan estas representaciones mentales?, ¿qué propiedad es la de tener “condiciones de verdad”? Esta pregunta incluye dos preguntas subsidiarias interrelacionadas:

⁷⁴ Una diferencia importante –sugerida por Millikan (1984, 1993)-, entre las representaciones de organismos no humanos y las representaciones mentales características de la psicología intencional humana, es que las primeras tienen al mismo tiempo la función indicativa y la imperativa. Por ejemplo, en el caso de la abeja, la danza *indica* la localización del néctar y al mismo tiempo *ordena* a las abejas expectantes a dirigirse hacia el néctar. La idea es que en el contexto de la cognición humana, en contraste, la función indicativa y la imperativa están desacopladas: las representaciones indicativas son creencias mientras los deseos son representaciones imperativas. Millikan (1993, 2004) y Sterelny (2003) han insistido que la separación de ambas funciones de representación es una importante transición en la evolución de la cognición. Otra diferencia importante es que las representaciones mentales humanas, a diferencia de las de otros organismos, no están directa o inmediatamente vinculadas con la producción de una acción consecuente.

⁷⁵ Hay por lo menos tres teorías clásicas al respecto (ver Cummins, 1989 y Millikan, 1993). La teoría *figurativa* o pictórica afirma que la representación es una relación de *semejanza*, es decir, que las representaciones son *como* lo que representan. La versión clásica del empirismo recurría a *imágenes* mentales, mientras que en la actualidad existe una versión matemática que recurre a funciones de *mapeo*. La teoría *causal / informacional* afirma que las representaciones son *acerca* de lo que las causan típica u originalmente. La teoría *simbólica* está inspirada en la idea de sistemas puramente formales o sintácticos que consisten en caracteres físicos manipulados por reglas que operan sobre su forma física. Bajo este esquema, una representación representa en virtud de su estructura interna dentro de un sistema de otras unidades igualmente estructuradas. Así, los símbolos pueden ser insumos y salidas de computaciones. Millikan (1984) defiende una versión matemática de la teoría pictórica en donde las representaciones mentales son concebidas como mapas. Revisaré esta teoría en la parte (b) del apartado tres de este capítulo.

- La pregunta por las *condiciones de corrección* de las representaciones. ¿Cómo pueden ser falsas estas representaciones mentales (los hechos o propiedades naturales que subyacen a la normatividad semántica)?
- La pregunta por la *individuación* de las condiciones de corrección. ¿Qué determina que una representación mental tenga un contenido particular en lugar de otro?

La respuesta que se da a 1 es, en último término, una cuestión empírica, y bien podría ser que los sistemas cognitivos dispongan de múltiples vehículos de representación. Puesto que el tema de este trabajo es la naturalización de la intencionalidad, el problema que me ocupa es 2 y no 1. Para estos fines podemos dar una respuesta a ‘2’ independientemente de la respuesta que demos a ‘1’. Sin embargo, si bien las teorías teleonómicas del contenido mental son en parte independientes de la respuesta que se da a la primera pregunta, en el tercer apartado presentaré la respuesta que ofrece Millikan a ese respecto.

2. El argumento teleosemántico.

El propósito de una teoría naturalista del contenido mental es la de caracterizar a las propiedades semánticas relevantes –las condiciones de verdad o de corrección de una representación mental–, en términos de alguna o algunas propiedades naturales no semánticas o intencionales. En otras palabras, la idea es mostrar cómo las capacidades representacionales de los estados mentales intencionales forman parte del mundo descrito por las ciencias naturales. Para ello es necesario satisfacer el criterio naturalista de adecuación (ver Fodor, 1984):

(CNA): *R* representa que *P* es el caso si y sólo si cierta relación *C* se da entre *R* y *P* independientemente de que *P* sea el caso o no, y donde *C* no puede a su vez estar caracterizada en términos semánticos o intencionales.

Lo que distingue a la propuesta teleosemántica de otros enfoques naturalistas es que se centra en torno a la noción de función teleológica o etiológica empleada en la biología de la

evolución revisada en el primer capítulo. La propuesta es que la teoría etiológica aplicada al contenido semántico de las representaciones mentales como las creencias y los deseos, pero también representaciones mentales menos sofisticadas, ofrece *prima facie* un buen argumento para su naturalización. En términos generales, la teleosemántica explica las condiciones de verdad de las creencias o las condiciones de satisfacción de los deseos en términos de las funciones biológicas de estos estados y de los mecanismos que les subyacen. ¿Cómo procede esta aplicación?

De acuerdo a la teoría etiológica revisada en el primer capítulo, la función propia de un rasgo biológico X de un organismo O es la de hacer Y si y sólo si

Definición (df): un caso actual de un rasgo X de un organismo O (miembro de una FRE) en un entorno E tiene la función de producir un efecto del tipo Y , si y sólo si, casos pasados (recientes) de X contribuyeron a la aptitud de los ancestros de O en E al producir Y y fueron seleccionados *por* hacer esta contribución a la aptitud de los ancestros de O frente a variaciones X en E .

Lo que podemos llamar la “inferencia teleosemántica” es la que aplica esta forma de explicación al contexto (psico)semántico para generar la siguiente hipótesis:

(HTS): El contenido de una representación caso R producida por un mecanismo cognitivo evolucionado M de un organismo O bajo condiciones C en el entorno E es la de representar C , puesto que representar C causó la selección de M en los ancestros de O en E , en virtud de que representar C mediante R tuvo un efecto positivo en la aptitud de los ancestros de O en E .

De acuerdo al esquema etiológico, M existe por el efecto seleccionado C ; así, la función propia de M es hacer (representar) C . La teleosemántica, en su sentido más modesto, se aplica a los rasgos fenotípicos cognitivos de los organismos como son los mecanismos de “detección”.⁷⁶ Sin embargo, la idea más ambiciosa es que recurrir a funciones propias para caracterizar estos mecanismos permite ir ascendiendo en grados de complejidad hasta poder ofrecer una versión teleofuncional de capacidades perceptuales y cognitivas más complejas como las que subyacen a

⁷⁶ Básicamente, se trata de los mecanismos cognitivos de los organismos que median un *respuesta adaptativa* específica con un elemento o elementos de su entorno al *registrar una señal ambiental específica* que le informa al organismos de la presencia de ese elemento o elementos ecológicamente relevantes (Ver, Millikan, 1984, Godfrey-Smith, 1991, 1992, 1996, Sterelny, 2003).

las actitudes proposicionales. Si pensamos a O como un agente intencional x , a R como actitud proposicional y a C como la proposición que P , de acuerdo a la propuesta teleosemántica tenemos que:

x cree que P si y sólo si x está en un estado cognitivo R que fue *seleccionado* por representar P , puesto que representar $P\dots$ (df).

x desea que P si y sólo si x está en un estado cognitivo R que fue *seleccionado* por realizar P , puesto que realizar $P\dots$ (df).

Esta formulación parece violar el CNA al apelar a la noción semántica de “representar”. Sin embargo, una teoría teleosemántica *pura* (Millikan, 2004) explica en qué consiste que una representación mental tenga contenido, esto es, de qué dependen las condiciones de corrección de las representaciones mentales, más no, *ipso facto*, cómo son ejecutadas estas funciones. De hecho, distintas versiones de la teleosemántica explican la ejecución de éstas funciones cognitivas mediante distintos mecanismos de representación:

x cree que P si y sólo si x está en un estado cognitivo R tal que fue seleccionado por *co-variar* con P (Papineau, 1993).

x cree que P si y sólo si x está en un estado cognitivo R que fue seleccionado por *mapear* el estado de cosas P (Millikan, 1984).

x cree que P si y sólo si está en un estado cognitivo R que fue seleccionado por *indicar* la presencia de P (Dretske, 1983).

Lo importante es que independientemente de qué teoríaelijamos sobre cómo son ejecutadas las funciones cognitivas (analizaré la versión de Millikan en el siguiente apartado), son las funciones propias o los efectos seleccionados de las representaciones mentales las que determinan el contenido de la representación, esto es, sus condiciones de verdad o de satisfacción.⁷⁷

⁷⁷ La postura de Millikan es ambigua a este respecto (ver 1984 y 2004). Ella considera que una teoría naturalista debe de separar una teoría de la representación verdadera de una teoría de la representación falsa, donde es la primera teoría la que da las condiciones de verdad. En este sentido, para ella la intencionalidad no es un solo fenómeno sino dos. Puesto que la primera teoría se explica en términos de relaciones de mapeo y la segunda en términos de funciones propias o de lo que es biológicamente normal no queda claro que sean las funciones propias de las representaciones las que dan el contenido, esto es, las condiciones de verdad. De cualquier manera no tenemos

La intuición general de estas propuestas es la siguiente: hemos visto en el primer capítulo cómo explicar el “propósito” o “función” de un rasgo natural sin apelar a un proceso intencional que presuponga la representación de la función. Ahora bien, en el contexto mental, sin embargo, lo característico del propósito psicológico que racionaliza las acciones de los agentes es que *presupone* la intencionalidad, esto es, presupone la representación mental de las metas o fines específicos que definen el “objeto” o la “dirección” del propósito. Decir que “presupone intencionalidad” implica que lo que el agente se representa son las condiciones de realización del objeto de su propósito.⁷⁸ ¿Cómo hacer de este “propósito representado” un caso de función biológica? Intuitivamente, el argumento es el siguiente: dado que según la teoría etiológica las asignaciones funcionales recogen de forma reductiva el elemento teleológico que subyace a tales asignaciones, y dado que en el contexto psicológico la teleología presupone intencionalidad, entonces, en principio es posible ofrecer el mismo tratamiento para el fenómeno de la intencionalidad, de tal suerte que ésta sea un miembro de estas categorías biofuncionales, heredando así, su estatus naturalista.

La estrategia de las teorías teleonómicas del contenido mental es entonces la de pasar de hechos históricos acerca de lo que causó la selección de un rasgo cognitivo particular a hechos constitutivos sobre lo que un rasgo cognitivo representa. En otras palabras, la idea es la de derivar el contenido de una representación mental a partir de la historia evolutiva de los mecanismos cognitivos que le subyacen y de los organismos portadores de dichos mecanismos. La motivación para identificar el contenido de una representación mental con la función propia de estos mecanismos es que la normatividad que acompaña la idea misma de contenido se explica en términos de la distinción entre lo que un rasgo tiene la *disposición* a hacer y lo que se supone que *debería* hacer. La manera de implementar esta estrategia es conceptualizando a los sistemas o mecanismos cognitivos y perceptuales a los que atribuimos representaciones mentales como mecanismos *evolucionados* –enfaticando por ende el carácter constitutivamente histórico de tales

que entrar en este problema exegético aquí. La idea de este apartado es que una teoría teleonómica del contenido identifica a éste con la función propia de la representación y con las funciones de los mecanismos que la producen y consumen. Más sobre esto en el tercer apartado.

⁷⁸ Por ejemplo, la acción de extender el brazo hacia el hueso de una quijada de burro en el suelo se podría explicar por la *intención* de Caín de matar a Abel. Esta intención a su vez se explica porque Caín *desea* que su hermano muera y *crea* que con esa quijada puede golpearlo de muerte. El agente actúa *para* lograr la realización de su propósito, deseo o intención guiado *por* la representación de éste, de suerte que su acción es *intencional*. En este caso, parece que el propósito es *causa* de la acción, al mismo tiempo que determina el *objeto* de la misma.

mecanismos- mediante la postulación de funciones selectivas, tanto a dichos *sistemas* cognitivo-perceptuales, como a los *estados* cognitivo-perceptuales particulares que generan. La teleosemántica propone entonces caracterizar al “contenido intencional” en términos de “función intencional” para de esta manera hacer de la “función intencional” un caso de “función evolutiva” (Walsh, 2002) reduciendo así las propiedades semánticas e intencionales a propiedades selectivas. Podemos ilustrar con un ejemplo la intuición fundamental del argumento.

Lo que un sujeto ve depende en parte de aquello a lo que el sujeto puede responder causalmente. Así, el contenido de un estado perceptual podría ser identificado con aquello que causa al estado. El problema es decir qué elemento en la cadena causal que termina con un estado perceptual debe ser designado como aquello que el sujeto está percibiendo, esto es, como el contenido de su percepción. Este es el llamado *problema de la profundidad* (Braddon-Mitchell y Jackson, 1996, 2007). Pensemos ahora en un percepto auditivo. Cuando oímos algo, nuestra experiencia auditiva posee un contenido determinado. Por ejemplo, oímos *el sonido de un balazo*. Desde una perspectiva causal de la percepción es el sonido el que causa el estado perceptual y por tanto lo que determina el contenido del estado. Sin embargo, el esquema meramente causal no parece ofrecer el contenido determinado que escuchamos -el del *sonido de un balazo*-, ya que tanto jalar el gatillo como la vibración al interior del oído constituyen igualmente causas activas que producen que tengamos el estado intencional pertinente. Sin embargo, lo que oímos no es ni el gatillo siendo oprimido ni la vibración al interior del oído; oímos el balazo. ¿Cómo determinar el elemento intencional dentro de esta cadena causal?

La propuesta teleosemántica es que la forma de determinar el contenido en este caso conlleva, primero, un análisis naturalista de la teleofuncionalidad de los mecanismos responsables de la percepción, en particular los auditivos, y segundo, la atribución al sistema auditivo de la función de transmitir la información acerca de la onda sonora dentro de cierto rango de frecuencia en el medio ambiente equivalente al sonido de un balazo. La idea es que oír la información que traducimos como “sonido de una balazo” y no la información del gatillo siendo oprimido o lo que ocurre al interior de nuestro oído hizo una diferencia en la aptitud de nuestros ancestros o de los ancestros de organismos con tales sistemas, tal que dichos mecanismos auditivos fueron seleccionados *por* arrojar tal contenido y no otro. Así, el contenido es *derivado* de una explicación a partir de *causas*, en particular de una explicación histórico-causal, pero es

individuado a partir de ciertos *efectos* históricos. La idea es que si podemos ofrecer éste análisis para la percepción, entonces hemos dado un primer paso para ofrecer el mismo tratamiento para el caso de las creencias y los deseos.

Sin embargo, hay razones para ser escéptico sobre la viabilidad de incluir a las actitudes proposicionales dentro del esquema teleosemántico. Principalmente destaca el hecho de que el contenido de muchas de esas actitudes nada tiene que ver con factores ecológicamente relevantes en el sentido estricto de incrementar la aptitud y menos con hechos relativos al pasado remoto evolutivamente relevante. De hecho, muchos de nuestros deseos claramente afectan negativamente nuestra aptitud. Por ejemplo, tenemos frecuentemente la expectativa de no estar embarazados o de no embarazar a nadie, deseamos fumar otro cigarro, pensamos en tomar un avión o creemos que los positrones son partículas sub-atómicas, etc. Una manera en que el teleosemanticista puede responder a esto es diciendo que los mecanismos relevantes que subyacen a la funcionalidad no tienen que ser necesariamente genéticos ya que algunos rasgos fenotípicos son resultado *además* de la influencia del medio durante la ontogénesis y, procesos ontogénicos como el aprendizaje, pueden incluir mecanismos de selección. Esto es posible puesto que, desde la perspectiva de la teoría etiológica, lo que hace a un efecto ser una función propia es independiente de si le subyace un proceso genético o no, pues el factor relevante es la intervención de un mecanismo de *selección* que descarte algunos efectos y promueva otros en virtud de algún *beneficio* que tiene para el objeto teleológicamente caracterizado. La idea es que el aprendizaje involucra mecanismos de selección que permiten adquirir funciones durante el desarrollo del organismo. Y esto tampoco nos compromete con una distinción categórica entre genético y adquirido, ya que los mecanismos de aprendizaje también han sido moldeados por la selección natural, lo cual quiere decir que hay una base genética para las distintas rutas de aprendizaje dados ciertos *entradas* ambientales. Así, la selección filogenética y la ontogénica se complementan, haciendo posible adquirir funciones y por ende representaciones sobre cosas totalmente nuevas en relación al origen histórico de tales capacidades. De la misma manera en que nuestro sistema digestivo no evolucionó para digerir comida chatarra y refrescos, nuestro sistema cognitivo no evolucionó para pensar en la postmodernidad o en quarks. Sin embargo, es necesario desarrollar una versión teleosemántica ontogénica a detalle, pues la mayoría de

nuestras representaciones son sobre aspectos relevantes durante el desarrollo. Este proyecto cae fuera del alcance de esta tesis.

3. ¿Qué son y cómo representan las representaciones mentales?

a) Las funciones propias de la cognición.

Los órganos como el corazón o los riñones poseen funciones propias “directas” (Millikan, 1984). Estas funciones son “relacionales” en el sentido amplio de que su efecto en la manera en que el organismo se enfrenta al entorno es determinante. Pero no son el único tipo de funciones propias directas. Más allá de rasgos fisiológicos, los rasgos y capacidades cognitivo-perceptuales de los organismos que derivan en conductas poseen funciones propias directas *relacionales* (Millikan, 1984). Éstas, si bien son “relacionales” en el sentido amplio arriba mencionado, lo son también en otro sentido más específico: establecen una “relación estructural abstracta” (Millikan, 2002) que supone la alteración de la relación entre el organismo y el medio, tal y como es necesario para que el ambiente provea de condiciones favorables para el organismo. La idea es que los sistemas o mecanismos responsables de estos cambios tienen una función propia de hacer que entre el organismo y el medio se dé una cierta relación *específica* entre ellos (Millikan 2002), esto es, que algún aspecto del organismo se corresponda con algún aspecto particular del entorno.

Por ejemplo, los mecanismos de detección de la rana poseen la función de detectar ciertos objetos en el entorno, presumiblemente moscas, tal que si se da esa condición en el medio, la representación de la rana del objeto relevante causa la captura de dicho objeto. El camaleón es sensible a cambios particulares en su medio que le permiten reaccionar de maneras apropiadas a dichos cambios adecuando el color de su piel al del entorno (camuflaje).⁷⁹ Lo que hace a estos ejemplos ser casos de función propia relacional es que ciertos estados internos tienen la función de corresponder o “estar dirigidos a” aspectos específicos del entorno. En contraste, el corazón o el estómago no tienen ninguna función propia *relacional*, en el sentido ahora introducido, en la

⁷⁹ La diferencia entre esos ejemplos es que en el caso del camaleón no hay una representación interna que medie entre el organismo y el aspecto relevante del entorno. Sin embargo, el cambio de piel relaciona de tal y cual manera específica al camaleón con un aspecto relevante del medio ambiente vinculado con su capacidad para sobrevivir y reproducirse.

medida en que no están “dirigidos a” ningún aspecto específico del entorno del organismo. En este sentido, estos órganos no tienen contenido; no representan nada.

Para Millikan (1984), los estados cognitivos representacionales tienen la función relacional de adaptar el progreso del sistema cognitivo a condiciones cambiantes en el entorno. Así, dichos mecanismos no sólo constituyen *históricamente* una adaptación al medio, sino que poseen una plasticidad adaptativa que corresponde a una “función propia adaptada” similar a la que encontramos en el aprendizaje y que relaciona sus productos particulares, a saber, creencias individuales, al medio dadas sus condiciones *actuales*. Dicha adaptabilidad no depende *directamente* de la selección natural, pero constituye una función propia *derivada* de la función propia relacional directa de los mecanismos que sí son resultado de la selección natural.

Esquemáticamente, si de hecho hay algo *X* –un ave en su nido- y otra cosa *Z* –digamos un depredador volando por encima del nido- que está situado de tal suerte que mantiene una relación con *X*, causando que *X* produzca *Y* –un canto característico que avisa del peligro-, entonces *Y* está “adaptada” en ese tiempo y en ese lugar a *Z* y, así, los mecanismos en *X* responsables de producir *Y* adquieren una función propia *adaptada*, función propia que es *derivada* de la función propia relacional directa del mecanismo que produce *Y* *más* el contexto actual. Una función propia “adaptada” derivada de una “relacional” directa se distingue entonces porque involucra el contexto actual. De ahí que sea posible “relacionarse con” o “representarse” cosas que no necesariamente estaban presentes en el episodio de selección pasada en la que los mecanismos pertinentes evolucionaron.⁸⁰ En este sentido, la función propia relacional de *Y* es indicar la presencia de un depredador, pero su función propia adaptada es indicar la presencia de *tal-y-cual-halcón-volando-aquí-y-ahora*. En resumen, la función propia relacional es de los mecanismos en *X* que deben relacionar a *X* con aspectos relevantes del entorno, mientras que la función propia adaptada producida por tales mecanismos es la situación de *Y* a un tiempo y lugar específico.

Esa distinción entre la función propia *relacional* directa de un mecanismo y su función propia *adaptada* derivada de la directa, es central para entender la diferencia entre los *mecanismos cognitivos* de formación de representaciones (o para el caso humano, de creencias) y

⁸⁰ Para Millikan (1984, 1993), las funciones adaptadas derivadas son suficientes para sortear la objeción de que la mayoría de nuestras representaciones mentales son sobre cosas completamente ajenas a las cosas que causaron la selección de los mecanismos cognitivos que le subyacen.

los *estados cognitivos* particulares (o creencia particular en el caso humano). Ambas cosas poseen funciones propias pero no del mismo tipo, aunque las segundas sean derivadas de las primeras. Otra forma de presentar esta distinción entre un mecanismo y estado cognitivo es si pensamos que las instancias de estados cognitivos de un organismo son causadas por sucesos que ocurren en el entorno del organismo, mientras que los mecanismos cognitivos, típicamente neuronales, son aquellos que *median* dichas transacciones causales. En este sentido, los estados cognitivos particulares *per se* no tienen funciones propias directas, como los mecanismos que las producen, sino que tienen funciones propias de forma derivada de las directas.

b) Mapas o íconos intencionales.

De acuerdo con Peirce, un signo o representación icónica es una relación *triádica* que involucra un vehículo de representación, lo representado y un intérprete. La idea es que entre el vehículo y el objeto de la representación existe algún tipo de *similitud* que el intérprete puede reconocer y que hace posible que el primero represente al segundo.⁸¹ Millikan (1984) utiliza una versión modificada de la caracterización de Peirce adecuada a una teoría naturalista de la intencionalidad: en lugar de hablar de un “intérprete” en el sentido de un agente que comprende las representaciones o signos que emplea, Millikan habla de un “consumidor” o de un mecanismo que “consume” la representación.⁸² Por otra parte, la “similitud” relevante en este contexto es una similitud *abstracta* que define como una función matemática de mapeo que establece, mediante una regla de proyección, un isomorfismo abstracto uno-a-uno entre aspectos de la representación o mapa y aspectos de lo representado o mapeado. La idea es que esta relación abstracta permite establecer una correspondencia entre transformaciones de ciertos aspectos de lo mapeado con

⁸¹ La relación de representación puede ser *diádica*. Algo X representa que P es el caso si y sólo si X es similar a Y o X covaría con Y o X es causada por Y, O X lleva la información que P. O *triádica*. Algo X representa que P es caso en caso de que y sólo en el caso de que hay algo Z (un agente, una comunidad de agentes) que usa X para representar Y. En éste caso, la relación entre la representación (vehículo) X y lo que es representado, P, no agota el contenido de la representación. La relación de representación, no se agota en lo representado, hace falta el modo de representar lo representado.

⁸² Originalmente (1984) Millkan hablaba de mecanismo de *producción e interpretación*. Pero puesto que no es necesario que el mecanismo intérprete entienda “lo que el signo signa” (Millikan, 1984, p. 96), posteriormente cambió la terminología a hablar de *consumidor*.

transformaciones de los aspectos correlativos del mapa.⁸³ Así, la relación triádica es entre el mapa, lo mapeado y un sistema o mecanismo que consume la relación mapa-mapeado. Millikan llama a estos mapas “íconos intencionales” y considera que éstos son la forma más simple y ubicua de intencionalidad en el reino animal.⁸⁴ Millikan (1984) ofrece la siguiente caracterización de un ícono intencional:

1. Un ícono intencional es un miembro de una familia reproductiva establecida y como tal posee funciones propias.
2. Una condición normal para la ejecución propia de la función de un ícono intencional es estar entre un mecanismo que lo *produce* y otro que lo *consume*, tal que ambos mecanismos tienen la función propia de *cooperar* y ésta es una condición normal para su operación propia.
3. Una función propia del ícono intencional es la de *adaptar al mecanismo consumidor a condiciones normales en el entorno* para que el mecanismo consumidor pueda ejecutar sus funciones propias bajo tales condiciones. Mientras los mecanismos responsables de producir y consumir los íconos intencionales constituyen adaptaciones históricas y, por tanto, poseen lo que Millikan (1984) llama funciones propias *directas relacionales*, los estados representacionales o íconos particulares tiene funciones propias derivadas de las directas que están *relacionalmente adaptados* a algún aspecto actual, no pasado, del entorno.⁸⁵

⁸³ La idea es que entre la representación o mapa y las condiciones mapeadas, la variación isomórfica se da de acuerdo a la *forma* de la representación o mapa, donde la forma está determinada de acuerdo con reglas de correspondencia específicas que constituyen la semántica del sistema representacional, donde estas variaciones o transformaciones exhiben una articulación entre aspectos variantes e invariantes de ambos términos de la relación.

⁸⁴ Los íconos intencionales pueden ser internos al organismo, como en el caso de la rana o externos al organismo, como en el caso de la danza de la abeja. Cuando el ícono es interno, los mecanismos de producción y de consumo están en el mismo organismo mientras que, si el ícono es externo, estos mecanismos se dan en distintos organismos.

⁸⁵ Mientras el concepto de *adaptación* es histórico, el concepto de *adaptatividad* no lo es. Un rasgo puede ser adaptativo porque produce un efecto positivo, en términos de aptitud, para el organismo portador, sin ser por ello una adaptación. Para ser una adaptación es necesario que el rasgo haya sido originado por selección natural o bien mantenido por selección natural. Ahora bien, si la adaptatividad de un rasgo más la suerte de una mutación genética causa la selección del rasgo por el efecto adaptativo, entonces el rasgo se convierte en una adaptación. Para hablar de adaptación la selección (y la herencia concomitante) es necesaria. Ahora bien, la idea detrás de la noción millikaniana de función propia relacional o función propia derivada es que algunas adaptaciones, como los mecanismos cognitivos o los mecanismos de camuflaje, consisten en generar estados adaptados a condiciones

4. En el caso de un ícono intencional *imperativo*, es una función propia del mecanismo consumidor, en la medida en que está adaptado al ícono, la de producir algo a lo que el ícono mapea de acuerdo con una función matemática de mapeo.
5. En el caso de un ícono intencional *indicativo*, la explicación normal de cómo el ícono adapta al mecanismo consumidor para que pueda ejecutar sus funciones propias hace referencia al hecho de que el ícono mapea con algo de acuerdo con una función matemática de mapeo específica.

Llamemos a los mecanismos de producción *MP* y a los de consumo *MC*. *MP* corresponde los mecanismos sensoriales y otros mecanismos cerebrales que dan lugar a representaciones cognitivas, mientras que *MC* corresponde a los mecanismos que emplean dichas representaciones para dirigir la conducta con el fin de realizar alguna función (“propósito”) biológico. Para Millikan, *MC* determina tanto el estatus intencional de una representación como su contenido. Puesto que las funciones son efectos que un rasgo se supone que debe producir, la función de una representación mental radica en la manera en que dicha representación contribuye a la función biológica de *MC*. La función de la representación permite que *MC* alcance su función biológica conectando a la conducta con las circunstancias representadas por la representación. Así, las condiciones de verdad o de satisfacción de una representación mental son las circunstancias bajo las cuales la conducta generada realiza la función *MC*. Una consecuencia de esta propuesta es que la determinación del contenido, de las condiciones de verdad o satisfacción, no son las condiciones de entrada, *MP*, lo cual permite dar cuenta de la normatividad de la intencionalidad (como veremos en detalle en el siguiente apartado): puesto que el contenido, las condiciones de verdad o satisfacción dependen de la salida conductual, el contenido es la condición bajo la cual el comportamiento resultantes *sería* apropiada, sean o no las circunstancias actuales que causan la representación de ese tipo.

Estas condiciones describen lo que un ícono intencional *es*, a saber, una función matemática de mapeo (iconocidad) más una función biológica cognitiva (intencionalidad), más no *de qué* es un ícono intencional (Millikan, 1984, p. 100). Aquello a lo que el ícono intencional

actuales cambiantes. En este sentido, esta plasticidad adaptativa puede ser considerada una adaptación (ver Sterelny, 2003).

está dirigido o aquello que mapea lo llama Millikan su “valor real”. Un ícono intencional posee un valor real cuando mapea aquello que tiene por función propia mapear. Si un ícono intencional no posee un valor real entonces el ícono no es acerca de nada o no está dirigido a nada y por tanto no ejecuta su función propia. Sólo si tal correspondencia se da con un valor real es que la parte consumidora de la representación (los castores expectantes por ejemplo), puede ejecutar su función propia que supone que dado tal contenido –dado tal lugar y tiempo- deben reaccionar de manera apropiada de acuerdo a una explicación normal: huir del peligro. De nuevo, tal explicación normal depende de una condición normal fundamental, a saber, la correspondencia entre el mapa o representación y la condición en el mundo mapeada, correspondencia que determina la forma de la variación. Por otra parte, en el viejo ejemplo de la abeja danzante, decimos que la parte consumidora del mecanismo intencional, las abejas expectantes, no ejecutaría su función propia de ir en la dirección específica del néctar de acuerdo a una explicación normal a menos que dicha localización corresponda de hecho con la danza. Esto es, que entre la danza y el lugar del néctar hubiera una correlación de transformación como danza-en-un-lugar-y-un-tiempo-con-tal-y-cual-orientación (Millikan, 1993, p. 91).

Ahora bien, ¿por qué Millikan recurre a una teoría figurativa o icónica de la representación como lo es la idea de “mapas” en la que lo que funda la relación de representación es una especie de *semejanza* entre la representación y lo representado? Es cierto que el modelo figurativo clásico de la representación se enfrenta a problemas serios que han inspirado otros modelos para dar cuenta de la relación de representación como es el modelo simbólico. Sin embargo, la visión clásica de la teoría figurativa fracasa porque al no especificar un intérprete, usuario o consumidor que determina *qué* aspectos son relevantemente semejantes entre la representación y lo representado, queda indeterminado lo que tienen en común la figura con lo figurado (Millikan, 1993). La razón es que la mera función de mapeo, en el sentido matemático, no puede garantizar qué aspectos son similares entre la figura y lo figurado, ya que hay una infinidad de maneras actuales de llevar a cabo esta correspondencia. Para sortear esto, la relación matemática de mapeo tiene que estar respaldada por algún otro vínculo real, sea causal, histórico o ambos, que garantice qué es lo que funda la correspondencia, es decir, la similitud abstracta entre el mapa y lo mapeado (Millikan, 1984, p. 86-87). ¿Cuál es la regla de mapeo que debemos mencionar para caracterizar lo que la danza de la abeja mapea? La respuesta está en la historia

evolutiva de las abejas. (Millikan, 1993, p.78). Pensar la semántica desde la perspectiva evolutiva permite sortear estos defectos de la teoría clásica figurativa.⁸⁶

c) Creencias y deseos.

Ahora podemos formular de manera más específica las funciones propias de las creencias y deseos que da Millikan (1984, 1993). Un primer factor a reconocer es que las condiciones de identidad-tipo de las creencias y los deseos difieren en relación a sus condiciones de identidad-caso, esto es, a lo que las individúa. La función propia de los deseos es la de *ayudar a causar su propia realización*. Esto quiere decir que las condiciones de identidad-tipo de los deseos se obtienen de la misma manera que categorías como “riñón”, “ojo” o “corazón”, en donde tales órganos son nombrados haciendo referencia a sus funciones propias más obvias. Cuando un deseo es de hecho realizado de acuerdo a su función propia más obvia, entonces dicho deseo posee un valor real. Por otra parte, los deseos tienen la función propia subsidiaria de participar en inferencias prácticas junto con otras creencias y deseos.⁸⁷

Ahora bien, es un hecho que las personas a veces desean cosas que son simplemente irrealizables por lo que las condiciones normales para la realización de un deseo no tienen que ser necesariamente condiciones realizables (Millikan, 1993, p. 95), ni siquiera en promedio, sino sólo las suficientes veces. La razón es que es de esperar que los deseos relevantes hayan participado ocasionalmente en procesos que logren afectar su realización, en particular los deseos directamente vinculados con factores centrales a la supervivencia y reproducción. Además, el que

⁸⁶ Por otra parte, es cierto que la teleosemántica no está necesariamente comprometida con la teoría pictórica. Sin embargo, una razón para preferir este modelo en su versión matemática de mapas es que parece claro que la danza de la abeja representa la localización de néctar porque *figura* aquello que representa, en lugar de, por ejemplo, ser causado por él. Esto es obvio ya que la danza ocurre sin que la abeja danzante o las abejas expectantes estén en presencia del néctar tal como requeriría una teoría de co-variancia. En cuanto a un teoría de co-variancia causal histórica, el problema aquí es las condiciones de satisfacción de un representación corresponden a las condiciones de entrada. Esto tiene dos dificultades. Primero que vuelve a la falsedad y el error más difícil de explicar pues presupone que la mayoría de las veces las representaciones son verdaderas, pues son causadas por aquello que representan. Segundo, esto desvincula la evaluación de la representación de la acción que produce, siendo que desde una perspectiva evolutiva las representaciones desvinculadas de acciones consecuentes parecen muy tardías. Además, el beneficio de representar algo, en términos de incremento en la aptitud esta directamente vinculado a las acciones que se siguen de representarse el entorno de cierta manera.

⁸⁷ Así, la “inferencia práctica” resulta de la combinación entre creencias y deseos de maneras novedosas para dar lugar a intenciones y después a acciones correspondientes, pero también las creencias se combinan para generar otras creencias (Millikan, 1993, p. 99).

los deseos tengan la función propia subsidiaria de participar en inferencias prácticas junto con otras creencias, permite evaluar la viabilidad del deseo y la manera de llevar a cabo su realización.

Con respecto a las creencias, a diferencia de los deseos, éstas no son individuadas por sus funciones propias más obvias. Lo que hace que una creencia sea sobre P y otra creencia sobre Q no depende de ninguna función propia *de* las creencias mismas, sino de ciertas condiciones normales que se han de darse *en el mundo* si dichas funciones propias han de ser ejecutadas o realizadas de acuerdo con una explicación normal (Millikan, 1993, p. 71 y 189). Más exactamente, la creencia de que P depende de la condición normal mapeada, a saber, P. Así, a diferencia de los deseos, la explicación normal de la función-contenido de una creencia depende fundamentalmente de ciertas “condiciones normales” *externas* a la propia creencia y que constituyen aquello que la creencia representa. Igualmente, si hablamos de una creencia en particular, lo que la identifica o individua en términos de su contenido es la condición normal específica, el aspecto relevante en el entorno con la que mantiene un relación de mapeo. Esto permite a Millikan decir que en cierto sentido “todas las creencias tienen la misma función propia” (1993, p. 89, nota 5), pues no son éstas las que las individúan, sino la correspondencia con la condición normal en el mundo que mapean.

Las creencias también poseen funciones propias subsidiarias. Millikan identifica dos: participar en inferencias para ayudar a ejecutar la realización de los deseos y participar en inferencias para producir otras creencias verdaderas. Así, el esquema de Millikan recupera el esquema de la psicología popular. Ahora bien, como hemos dicho, estas funciones propias son subsidiarias, lo cual quiere decir que ninguna de ellas ofrece condiciones de identidad-tipo para las creencias. En este sentido, una creencia no es identificada diciendo qué deseos ayuda a realizar o qué otras creencias verdaderas puede producir mediante procesos inferenciales dadas otras creencias. Por ejemplo, si dos personas, *K* y *W*, tienen la creencia de que hay un fraude electoral, las “condiciones de identidad” bajo las cuales podemos decir que tienen la *misma* creencia no hace referencia ni a los *conceptos* con lo que *K* y *W* describirían o reconocerían el fraude o las elecciones, ni tampoco a las *inferencias* específicas que cada uno estaría en disposición de hacer a partir de qué otras creencias sostienen. Lo que tienen en común *K* y *W* es que cada uno posee una estructura interna (un mapa) tal que si y sólo si *de hecho hay un fraude electoral* es que se pueden realizar tanto la función de participar en inferencias para producir

nuevas creencias verdaderas, como para ayudar a realizar los deseos de forma normal. Pero sin esta condición normal en el mundo, decimos que sus creencias tan sólo “se supone” deben ejecutar tales y cuales funciones aunque de hecho no las realicen. En ausencia de tal condición normal decimos que la creencia es falsa y, por tanto, que no puede ejecutar sus funciones propias. Pero más aún, tal creencia es *defectuosa* precisamente porque en ausencia de dicha condición, los mecanismos que consumen las creencias no pueden ejecutar sus funciones propias. En este sentido, la falsedad no emana de una falla en los mecanismos de formación de creencias, pues tal anormalidad no está *en* el agente y en sus sistemas cognitivos, sino que son las condiciones *externas*, las del mundo, las que no son anormales para la ejecución adecuada de sus funciones. Esto quiere decir que las condiciones externas son diferentes de las condiciones bajo las cuales tales mecanismos evolucionaron y fueron seleccionados. En el siguiente apartado presentaré en detalle el modo en que esta propuesta teleonómica da cuenta de la normatividad semántica. Pero antes unas palabras más sobre la implicaciones de esta propuesta teleosemántica sobre la individuación de las representaciones mentales.

Toda teoría del contenido semántico, naturalista o no, debe pronunciarse ya sea por un *internismo* o por un *externismo*, o bien por alguna combinación de éstos. En términos generales, para el externismo semántico el contenido representacional de un estado intencional depende de cómo es el mundo fuera del agente al que atribuimos el estado. Por su parte, el internismo niega que lo que ocurre fuera de la cabeza del agente sea relevante para la individuación del contenido representacional de su estado intencional. Así, mientras el externismo concibe el contenido semántico en sentido *amplio* y, por tanto, como una propiedad *relacional*, el internismo lo concibe en sentido *estrecho* como una propiedad *local*, esto es, como una propiedad que sobreviene sólo en lo que ocurre en el sujeto al margen de los detalles de su entorno y de su historia. En relación a estas alternativas, podemos decir que la teleosemántica ofrece un fundamento biológico evolutivo para el externismo: si las condiciones de identidad de la función propia de un rasgo dependen de las condiciones del entorno ecológico y de la historia causal, y el contenido representacional de una creencia es un caso de función propia cognitiva, entonces éste es relacional y la relación relevante es histórica de cierto tipo (Millikan, 1993, p. 29). El fundamento naturalista para defender una concepción del contenido representacional en sentido amplio es entonces un modelo ecológico de la cognición.

4. ¿Cómo pueden ser falsas las representaciones mentales?

La posibilidad de error es parte de lo que queremos decir cuando describimos una capacidad como cognitiva o representacional. Esta normatividad genera el siguiente problema para el naturalismo; ¿cuál es la *fente* de esta normatividad? Si aceptamos que la mentalidad y, en particular, la intencionalidad o semánticidad, existe *independientemente* de nuestras prácticas interpretativas, entonces las normas no pueden provenir de las intenciones, propósitos y deseos humanos, ya que éstos dependen de la existencia previa de normas no humanas. Podemos empezar motivando el problema en el contexto de una formulación causal del contenido representacional (Fodor, 1990).

Supongamos que el símbolo “alacrán” significa *alacrán* y, por tanto, que sólo los alacranes caen bajo la extensión del símbolo en la medida en que es una regularidad (incluso quizá una ley) que los alacranes causen el símbolo “alacrán”. Ahora bien, también es cierto que algunas cosas que no son alacranes pueden causar y de hecho causan símbolos o representaciones de “alacrán”. Por ejemplo, una cucaracha-en-la-oscuridad puede ser suficiente para causar “alacrán”. Puesto que una representación en la teoría causal tiene por contenido aquello que la causa, el símbolo “alacrán” representa entonces la propiedad de ser alacrán-o-cucaracha-en-la-oscuridad, o lo que es lo mismo, la extensión de “alacrán” incluye tanto a los alacranes como a las cucarachas-en-la-oscuridad. La consecuencia indeseable es que las representaciones de “alacrán” que son causadas por cucarachas en condiciones de oscuridad resultan verdaderas. El problema es, entonces, que no podemos distinguir entre un caso de una representación verdadera que tiene por contenido algo disyuntivo y un caso de representación falsa que no posee un contenido disyuntivo.⁸⁸

⁸⁸ El “problema de la disyunción” (Fodor, 1990) es una versión del problema normativo para los enfoques de covariancia causal sobre la representación. Según éstos, los estados internos de un agente u organismo que típicamente covarían con la presencia de algo en su entorno son acerca de dicho aspecto del entorno. El problema es que si un estado interno covaría con vacas también lo hará con la presencia de cosas muy similares a vacas, así que la dificultad estriba en explicar cómo estos estados pueden tener el contenido ‘vaca’ y no ‘vaca u otro objeto similar’. Mientras el problema de la falsa representación (o normativo) vuelve al error algo imposible al *descartar* la representación de una situación cuando dicha situación no existe, el problema de la disyunción, por otro lado, hace al error imposible al *incluir* la representación de demasiadas situaciones.

La cuestión apremiante es la de ofrecer un criterio naturalista para trazar una distinción entre usos o condiciones que *fijan* el contenido de la representación y usos o condiciones que no lo hacen, en donde los usos o las condiciones que llevan al contenido disyuntivo deben ser separados de los que fijan al contenido, para hacer viable una explicación de la posibilidad del error y la falsedad. Tenemos entonces que hay circunstancias bajo las cuales las representaciones tienen aquello que representan como una condición o causa necesaria y/o suficiente de su producción. Tales condiciones fijan el contenido de la representación. Se sigue que las representaciones falsas responden a representaciones que son producidas bajo otras circunstancias. Podemos ahora preguntar, ¿cuáles propiedades naturales de las representaciones (mentales) definen las condiciones que fijan su contenido y que son asimétricas con respecto a las condiciones que dan lugar a la falsedad?

Ahora bien, independientemente de la teoría causal de la relación representacional o de cualquier otra teoría al respecto, la propuesta teleosemántica de Millikan es una manera de responder a la pregunta por la fuente de la normatividad apelando al proceso biológico de evolución por selección natural. La respuesta consiste en hacer la siguiente identificación: las *condiciones de corrección* de las representaciones mentales son *condiciones (históricamente) normales* de ejecución propia. Lo cual implica que las propiedades *semánticas* son propiedades *selectivas*, ergo, que las propiedades semánticas son históricas. Pero vayamos por partes.

Puesto que son las condiciones de corrección las que individúan a las representaciones, el contenido de una representación mental se identifica con la función propia de esa representación y de los mecanismos que la producen y o consumen. De esto se sigue que una representación que es empleada y / o producida bajo condiciones *anormales* no ejecuta su función propia y, sin embargo, retiene la función; así una representación mental retiene su contenido independientemente de ser falsa o vacía. Tal como una estructura o rasgo biológico puede tener una función y, sin embargo, no ejecutarla o ejecutarla incorrectamente, igualmente si construimos el contenido de una representación como la función propia de dicha representación, ésta puede poseer la función putativa y no obstante ejecutarla incorrectamente en ausencia de condiciones normales, o simplemente no ejecutarla por ausencia de las condiciones que disparen la ejecución o por malformación. En pocas palabras, la *falsedad* o el *error* cognitivo se explican como un caso

de *disfunción* en la ejecución de una función propia relacional en la que una representación está involucrada.

El argumento central es entonces que si caracterizamos al contenido intencional como miembro de una categoría teleofuncional, podemos dar cuenta del carácter normativo de la intencionalidad como derivado de la normatividad inherente a las asignaciones teleofuncionales. Así, de la misma manera en que la marca característica de la intencionalidad es que dichos estados pueden ser falsos o insatisfechos, esto es, “acerca” de lo que no es el caso, así también la marca característica de la teleofuncionalidad, de lo que tiene una función propia, es que puede no ejecutar la función o propósito putativo (Millikan, 1993). La intencionalidad, el contenido representacional bajo esta teoría, es entonces una relación histórica entre lo que el signo o representación debería significar o representar y lo que de hecho significa.

Supongamos que *E* es un estado cognitivo de un organismo *O* que tiene la función propia -derivada de la función propia del mecanismo cognitivo evolucionado *M*-, de representar *X* (alacranes), y sin embargo, en ocasiones *E* representa otra cosa, digamos *W* (cucarachas). ¿Cómo puede una representación representar algo que no está incluido dentro de su contenido-función? La respuesta teleosemántica supone la siguiente historia. *X* es un aspecto del entorno ecológico evolutivamente relevante para un organismo *O*. *E* y *M* fueron seleccionados porque representaban *X*, ya que hacerlo incrementó la aptitud de los ancestros de *O*. Así, a través de las generaciones *E* adquirió la función propia de representar *X*. Ahora bien, *X* produce *E* bajo condiciones (históricamente) normales, esto es, bajo las condiciones en las que los ancestros de *O* y en particular *M*, evolucionaron.⁸⁹ Pero en ocasiones *W* en lugar de *X* produce *E*. ¿Cuáles son estas “ocasiones”? La respuesta es que se trata de ocasiones anormales. Cuando *W* da lugar a *E*, *E* es una representación disfuncional o defectuosa: *E debería* representar *X* pero *de hecho* representa *W*. Así, sólo cuando *E* representa *X* es que *E* representa *intencionalmente X*.⁹⁰

Puesto que las condiciones relevantes que fijan el contenido son las condiciones históricamente normales de ejecución propia de los mecanismos representacionales, el

⁸⁹ En este caso “*O*” forma parte de una familia reproductiva establecida de primer orden mientras los mecanismos *M* que produce *E* forman parte de una familia reproductiva establecida de orden superior.

⁹⁰ Millikan (2002, p. 65) va más lejos al afirmar que podemos decir que una representación que fracasa en cumplir su función de representar simplemente no representa nada. Sin embargo, no tenemos que aceptar esta afirmación ya que, intuitivamente, aunque una representación sea falsa, aun así, si conocemos su contenido, podemos decir lo que se supone que debería haber representado.

teleosemanticista puede circunscribir la posibilidad de error a las normas biológicas inherentes a la historia causal (selectiva) de dichos mecanismos. Así, el *espacio lógico y natural* para la evaluación semántica de los estados cognitivos es un espacio histórico-selectivo. La idea central es que estas normas semánticas son genuinas normas naturales, en oposición a normas morales, éticas o sociales, porque establecen una medida –sin apelar a un agente o comunidad de agentes-, a partir de las cuales los casos actuales pueden concordar o divergir. Así, estas normas naturales son descriptivas y no prescriptivas.⁹¹

En resumen, atribuir una función propia es atribuir una propiedad normativa natural en tanto que permite discriminar las ejecuciones que concuerdan con la función de las que divergen de ella. Dicha medida es histórica ya que permite medir hechos presentes a partir de hechos pasados. Esta diferencia entre el pasado y el presente funda la asimetría entre lo que una estructura históricamente se supone *debería* hacer y lo que de *hecho* hace, lo cual supone que el criterio normativo –la función- y las ejecuciones actuales de dicha función, no están al mismo nivel. La propiedad normativa teleofuncional, si bien es una propiedad *real y natural* de la estructura o entidad biológica, no es una propiedad *actual o presente* de dicha estructura, porque la fuente de la normatividad y de la funcionalidad es la *historia* evolutiva de la estructura, rasgo o carácter en cuestión. Así, no todas las circunstancias *actuales* en las que una función putativa se ejecuta son necesariamente condiciones *normales*. Podemos contrastar las condiciones actuales y las condiciones ideales sólo si conocemos la historia evolutiva particular de la entidad en cuestión.

Una virtud de la resolución teleosemántica del problema normativo es que permite resolver la paradoja de Brentano (Millikan, 1984, 2002). El teleosemanticista puede decir que Brentano estaba equivocado en pensar que sólo algo mental puede mantener una relación con algo inexistente. La idea es que, al igual que una representación mental retiene su contenido independientemente de su valor de verdad, en particular independientemente de que su valor de verdad sea falso, un rasgo retiene su función propia independientemente de que pueda ejecutarlo

⁹¹ Además, como hemos visto ya en el primer capítulo, estas *condiciones normales* no son ‘normales’ en el sentido estadístico o promedio, sino en el sentido de condiciones *ideales* o históricamente *óptimas*. Así, mecanismos que rara vez ejecutan sus funciones propias o que la mayoría de las veces las ejecutan disfuncionalmente, aun así retienen sus funciones propias, siempre y cuando sus ancestros las hayan ejecutado las suficientes veces para hacer una diferencia en términos de la supervivencia y reproducción tanto de dichos mecanismos como de sus portadores.

o de si lo ejecuta propiamente. La relación intencional, *viz a viz*. la posibilidad de que uno de los *relatum* no exista, consiste en una relación natural históricamente mediada, en la que el primer *relatum* puede darse sin que se dé el segundo en virtud de que, en el pasado selectivamente relevante, ambos *relata* se dieron de hecho y dicha correlación causó la selección. La intuición es que el hecho de que una función corresponda a un rasgo a pesar de que éste no posea las disposiciones adecuadas para ejecutarlo, no nos inclina a decir que la función posea algún tipo de “inexistencia intencional”.⁹² En cambio, es en virtud de la historia de los ancestros del rasgo que podemos afirmar que no obstante *tiene* propósito o función. La consecuencia relevante para el proyecto naturalista es que no tenemos que postular entidades ontológicamente sospechosas que medien entre el signo o representación y lo signado o representado. La única mediación real es la historia de los mecanismos con capacidades representacionales.

5. Conclusión: consecuencias metafísicas del naturalismo Darwinista.

“Life and mind have a common abstract pattern or set of Basic organizational properties. The functional properties characteristic of mind are an enriched version of the functional properties that are fundamental to life in general. Mind is literally life-like.”

Godfrey-Smith (1996).

La propuesta teleonómica del contenido mental satisface el requisito naturalista: las propiedades semánticas son *reducidas* a propiedades no-semánticas sino físico-naturales, en particular, a propiedades biofuncionales de rasgos evolucionados. Las propiedades semánticas de las representaciones mentales que determinan las condiciones de corrección de una representación mental son propiedades selectivas de dichas representaciones en virtud de la historia causal de

⁹² En el sentido general en que la intencionalidad supone una relación donde uno de los *relatum* no existe y donde el error es posible –la famosa normatividad–, todos los rasgos biológicos que poseen funciones propias involucran intencionalidad (Millikan, 1984). El corazón puede no bombear sangre o hacerlo de forma disfuncional y sin embargo retener su función. Esto es posible porque mientras la función propia de un rasgo es una *propiedad selectiva pasada*, la ejecución es una *propiedad disposicional* (presente). En este sentido amplio, todas las cosas con funciones propias tienen condiciones de corrección, esto es, condiciones de ejecución normal o propia en donde lo que de hecho hace un rasgo y lo que debería a ser depende de la asimetría entre la historia causal del rasgo y sus propiedades causales actuales. Sin embargo, sólo los rasgos que implementan funciones matemáticas de mapeo, rasgos cognitivos, tiene contenido intencional estrictamente, pues se corresponden con aspectos específicos del entorno. De ahí que los teleosemanticistas expliquen el error cognitivo y la representación falsa en este sentido. La normatividad es reducida a propiedades funcionales de los rasgos evolucionados. Esta reducción identifica las condiciones de identidad de las representaciones mentales con propiedades selectivas pasadas.

los mecanismos cognitivos y perceptivos que les subyacen: un *efecto pasado* – el hecho de que *R* representó *P* y que hacerlo tuvo un efecto positivo sobre la aptitud de los portadores de *R*-, explica la ocurrencia de ese efecto en el presente, esto es, que un caso de *R* se pueda dar sin que *P* se dé. Así, la explicación etiológica de cómo es posible que se dé *R* en ausencia de *P* es que en el pasado *R* y *P* fueron selectivamente correlacionados y un rasgo, en este caso *R*, retiene su función, representar *P*, aunque de hecho no la ejecute.⁹³

La originalidad de este naturalismo radica en que la reducción no es física sino biológica y, como tal, está muy lejos de *eliminar* las propiedades (mentales) de nivel superior. Creo que esta distinción es importante pues la diferencia entre una reducción física y una biológica es tanto una diferencia *epistémica* como una *metafísica*. Es epistémica pues supone que una psicología científica intencional generaría explicaciones en términos ecológico-evolutivos. En este sentido, la teleosemántica determina qué tipo de descripciones y modelos resultan relevantes para investigar empíricamente el fenómeno de la intencionalidad. En cuanto al aspecto metafísico, lo que tenemos es que las propiedades semánticas son de un tipo particular, a saber, propiedades histórico-causales y no propiedades disposicionales, (aunque sí naturales). Me concentraré en el elemento metafísico. ¿Cómo puede ser la intencionalidad una propiedad histórico-causal?

La idea es que el pensamiento y la percepción son intencionales en tanto representan cosas y las cosas que representan no tienen que existir para poder ser pensadas o tener percepciones acerca de ellas. Si asumimos que los pensamientos y las percepciones son estados cerebrales, la pregunta es, ¿qué hace a un estado cerebral representar una cosa y no otra o no representar nada?, ¿qué tipo de relación es ésta en la que uno de los *relata* es inexistente? La propuesta teleosemántica es que el estado cerebral no es ni una propiedad intrínseca del cerebro

⁹³ Es importante dejar claro en qué sentido estos “efectos pasados” son causalmente eficaces. Puesto que la selección natural no es un proceso intencional y por ende no anticipa sus resultados, los efectos seleccionados no nacen siendo funciones, sino que adquieren su estatus funcional sólo *retrospectivamente*. De ahí que la teoría etiológica histórica ofrezca un análisis reductivo de la teleológica de ciertos rasgos mediante la historia causal de los efectos de dichos rasgos. En otras palabras, es el carácter genealógico el que hace a los efectos pasados ser causalmente eficaces con respecto a los efectos del mismo tipo en el presente sin caer en los problemas de la teleología clásica pre-darwinista, en la que los efectos afectan a su vez las causas que los producen: los efectos no afectan sincrónicamente a sus causas, sino que explican en parte o son una de las causas de la presencia de ese rasgo en generaciones subsecuentes. Es entonces el elemento histórico el que permite respetar la idea de que las causas anteceden a los efectos, al tiempo que podemos asignarle un rol causal a los efectos en la producción diacrónica de sus causas en la siguiente generación.

ni una propiedad extrínseca relacional actual. Sólo las relaciones extrínsecas *pasadas* explican las propiedades intencionales de los estados cerebrales.

La motivación para explicar la intencionalidad en términos históricos puede ser presentada con un ejemplo. ¿Cómo es posible pensar que los nopales son verdes sin estar ante un nopal verde? Si no estoy en presencia de nopales verdes entonces entre mi pensamiento y el objeto de mi pensamiento no hay una relación real, esto es, una relación en la que ambos *relata* existen. La idea es entonces que la relación entre mi pensamiento y su objeto hay una relación histórica. Las propiedades intencionales de los estados cognitivos son resultado de relaciones reales que existieron en el pasado: en términos ontogenético diríamos que puesto que en el pasado he estado ante nopales verdes y éstos han causado ciertas percepciones y pensamientos, puedo ahora pensar en ellas en ausencia de ellas. En términos evolutivos (filogenéticos) la idea es la misma: puesto que mis ancestros estuvieron ante nopales verdes y percibirlos tuvo un efecto positivo sobre su aptitud...

Esquemáticamente podemos preguntar, ¿qué hace que R represente P ?, ¿Qué es lo que hace que entre R y P haya una conexión semántica que va de R a P ? La respuesta no puede ser que hay una relación real –en el sentido de una relación presente o actual- entre R y P ya que R puede tener el contenido P sin que P exista. En cambio, R o casos del tipo R pasados y casos de P en el pasado estuvieron en la relación real apropiada. Las propiedades intencionales y semánticas pueden ser fundadas en relaciones reales con objetos existentes en el pasado. Así, la misteriosa relación intencional es analizada en términos de relaciones reales pasadas, consiguiendo así su estatus naturalista.

Problema: apelar a relaciones históricas explica que pueda pensar acerca de cosas con las que no estoy relacionado en el presente pero con las que, al menos en principio, se pudo tener una relación en el pasado, ¿Qué hay de las cosas que puedo pensar con las que ni siquiera estuve relacionado en el pasado, esto es, que nunca, no ahora ni antes, han existido tales como unicornios o la Fuente de la Juventud? Quizá la respuesta radica en que esos casos son resultado de vincular mentalmente cosas o propiedades de cosas que sí existieron (imaginación). En cualquier caso, es necesario saber más sobre cómo operan la filogenia y la ontogenia en la generación de creencias y deseos particulares. En este sentido, la propuesta teleosemántica es apenas un modelo abstracto que requiere ser llenado con muchos detalles y donde el principal

desiderata es generar una teoría convincente de cómo se pueden producir efectos seleccionados durante el desarrollo.

¿Qué implicaciones tiene el recurso a la *historia* causal para el tipo de naturalismo que se sigue de aceptar la propuesta teleosemántica? Para terminar argumentaré brevemente en qué sentido esta reducción dista mucho de ser una “eliminación”, dando lugar a una versión integradora de lo mental a lo natural. La idea general es que, si bien la teleosemántica como propuesta naturalista *sí* implica en un sentido una reducción de propiedades psicológicas a biológicas –en particular de propiedades intencionales a propiedades biofuncionales-, el que la reducción sea a propiedades *biológicas* y no directamente a propiedades físicas permite hablar al mismo tiempo de una *integración* de las propiedades psicológicas al mundo natural.

Primero, porque tanto las propiedades biofuncionales como las mentales sobrevienen en propiedades físicas, esto es, las funciones evolutivas tienen múltiples realizadores físicos. Por ejemplo, la morfología de los corazones varía entre las especies, no obstante, todos los corazones tienen la función de bombear sangre. Esto es una consecuencia de que las explicaciones funcionales apelen a *efectos* (seleccionados) e indirectamente a causas: las propiedades biofuncionales son propiedades perfectamente naturales que no son idénticas a las propiedades disposicionales de las estructuras que las instancian.

Segundo, porque si hemos dicho que las características constitutivas de la intencionalidad son su carácter relacional y su carácter normativo, en la propuesta teleosemántica estas características *en sí mismas* no son reducidas ni a propiedades locales o no relacionales en el primer caso, ni a propiedades meramente causales o no normativas en el segundo caso.⁹⁴ En cambio, dichas propiedades ya existen como tales en el mundo natural, por lo que la naturalización procede identificando dichas propiedades mentales con esas mismas propiedades en el mundo natural biológico. En otras palabras, tales propiedades no son “otra cosa” (Fodor, 1987) a la que haya que reducirlas sino que son ya, de cabo a rabo, cosas naturales, pues no son

⁹⁴ Recordemos que en la introducción se dijo que la mera relación de *sobrevenida* de las propiedades mentales en las físicas no garantizaba su eficacia causal. Pero que en cambio, si las instanciaciones o realizaciones de las propiedades causales mentales tuvieran no sólo una relación de *dependencia* ontológica sino una relación de *constitución* con las instanciaciones de las propiedades físicas que sí son causalmente eficaces, entonces no habría problema en suponer que lo mental es causalmente eficaz, respetando así el principio de clausura causal de lo físico. El impedimento que señalamos para aceptar dicha relación de constitución era que la intencionalidad de lo mental involucra una propiedad relacional y normativa, mientras que las propiedades físicas, en particular las neuronales, son locales y ciertamente no son normativas.

propietarias de la mentalidad sino que el mundo biológico, en la medida en que exhibe funciones etiológicas, exhibe normatividad en el sentido robusto de Normalidad histórica. En este sentido, la intencionalidad *sí* puede sobrevenir en propiedades físicas *a pesar* de ser una propiedad relacional, de la misma manera que la función de un rasgo sobreviene en su estructura física y sin embargo incluye, de manera constitutiva, los detalles del entorno en su caracterización. Podemos decir que las propiedades biofuncionales son propiedades físicas en sentido amplio, o lo que es lo mismo, que no todas las propiedades físicas son intrínsecas.

Ahora bien, otro aspecto fundamental para el tipo de naturalismo aquí defendido es que las propiedades selectivas o biofuncionales a las que reducimos las propiedades semánticas también sobrevienen sobre propiedades físicas y, no obstante, son causalmente eficaces. ¿Qué las hace ser causalmente eficaces a pesar de ser sobrevenientes?

La razón es que las propiedades biofuncionales no sólo sobrevienen –en el sentido de dependencia–, sobre sus realizadores físicos, sino que guardan una relación ontológica de constitución con sus realizadores, y lo hacen en dos sentidos. Por una parte, porque la *forma* del rasgo constriñe a la *función*. Esto quiere decir que, si bien distintas estructuras o formas orgánicas pueden realizar la misma función, no cualquier estructura puede sin embargo realizar cualquier función, pues hay limitaciones estructurales al tipo de función que puede ejecutar dicha estructura. Por ejemplo, no cualquier estructura puede bombear sangre; hace falta, si bien mediante distintos recursos disposicionales y materiales, tener un mecanismo de bombeo. Por otra parte, la función constriñe a la forma. Esto se hace patente sobre todo en el caso de rasgos complejos donde las partes parecen acopladas direccionalmente para ejecutar una función que va siendo optimizada en la medida en que se van acumulando modificaciones en el tiempo, como lo muestra el caso paradigmático de la evolución del ojo. Lo que tenemos entonces es una relación de constitución en la medida en que la dependencia ontológica va en ambas direcciones: tanto de los realizadores físicos a las propiedades sobrevenientes (de la morfología del rasgo a su función), como de las propiedades sobrevenientes a los realizadores físicos (de la función del rasgo a su morfología). Y si la dependencia ontológica es bidireccional, entonces podemos hablar

de constitución ontológica y respetar el principio de clausura causal e incluir dentro de éste propiedades relacionales como siendo causalmente eficaces.⁹⁵

La importancia del enfoque teleosemántico radica entonces en hacer patente la relevancia de la biología –tanto evolutiva como del desarrollo–, para la mentalidad poniendo a disposición del filósofo una rica ontología, la cual, sin ofender las intuiciones fisicalistas, recupera la complejidad de la naturaleza viva y de uno de sus productos, la mente. Con ello exhibe el maniqueísmo de las formulaciones naturalistas típicas que hablan dicotómicamente de lo “mental” y lo “físico”, eludiendo así el papel mediador de la biología. Quizá no es más que la herencia cartesiana que llega hasta nosotros a través de Brentano la que genera la presión por formular maniqueamente el problema del lugar de la mente en el mundo natural. Pero fue Darwin quien sugiere una manera de superar los términos dualistas de la discusión, y la filosofía *no* puede pasar por alto esto. El mundo natural está compuesto tanto de un mundo meramente *físico* como de un mundo *vivo*, y es en éste donde la complejidad y la historia son relevantes. La mente y la intencionalidad no surgieron directamente de lo físico, por lo que debemos resistir la presión a explicarlas directamente a partir de lo físico. En cambio, nuestra *vida mental* y las de otras especies surgió de las contingencias acumuladas y acopladas, a través del tiempo, que forman el rico afluyente de la historia natural de los seres vivos, nosotros incluidos.

⁹⁵ La eficacia causal de estas propiedades funcionales emergentes se explica primero por su eficacia causal en relación a la aptitud, la cual es medible, ya que la ejecución Normal de una función causa un incremento en la aptitud mientras que su ejecución anormal causa una disminución en la capacidad del organismo portador para sobrevivir y reproducirse. Y segundo, si la función ya no es ejecutada, la eficacia causal de la emergente biofuncional se explica históricamente, esto es, por su eficacia causal pasada: por el efecto positivo sobre la aptitud de los ancestros del organismo actual que porta el rasgo biofuncional.

CAPÍTULO TRES:

La Objeción de la Indeterminación y su Disolución

1. Introducción.

En este capítulo revisaré una de las objeciones más importantes que se han formulado contra la teleosemántica, *viz a viz*, el problema de la indeterminación del contenido (Fodor, 1990, 1998).⁹⁶ En el capítulo anterior revisamos el modo en que la teleosemántica da cuenta de la característica normatividad de los estados representacionales. El problema que nos ocupa ahora tiene que ver con la característica *opacidad* de los contextos de atribución de estados intencionales: los predicados mentales de los enunciados que atribuyen estados intencionales no pueden sustituirse por predicados co-extensionales *salva veritate*. Dos representaciones mentales pueden ser co-extensionales o ser verdaderas bajo las mismas condiciones y sin embargo diferir en su contenido. Por ejemplo, *x* puede creer que *P* sin creer que *Q* incluso si ‘*P* si y sólo si *Q*’ es verdadera. Así, lo que determina las condiciones de corrección de una representación mental particular debe ser capaz de exhibir esta diferencia de contenido. El problema, dice la objeción, es que si construimos las condiciones de corrección como *condiciones (históricamente) normales de ejecución*, la diferencia no se respeta: el contenido de representaciones equivalentes resulta idéntico. De acuerdo a la objeción, la teleosemántica no es capaz de dar cuenta de esta opacidad porque los contextos de atribución funcional son transparentes.

Si bien el problema fue originalmente (Fodor, 1990) planteado como una objeción contra la teleosemántica, la disputa por cómo determinar el contenido de una representación mental es objeto de debate entre filósofos que rechazan la putativa indeterminación y que aceptan alguna versión de la teleosemántica.⁹⁷ En este, contexto el debate gira en torno a la manera en que la etiología de un mecanismo cognitivo determina el contenido de un estado cognitivo. En este capítulo argumentaré que, a) la putativa indeterminación no constituye una objeción contra el

⁹⁶ Ésta no es la única dificultad que ha sido planteada a la teleosemántica. Por cuestiones de espacio me concentraré en ella exclusivamente.

⁹⁷ Dennett (1988, 1996), en cambio, acepta la indeterminación del contenido pero considera que lejos de ser una objeción contra la teoría darwinista del contenido mental es una “precondición” para dar un recuento evolucionista de la representación mental (1996). No revisaré esta propuesta en este ensayo.

programa teleosemántico y b), hay buenas razones para elegir un recuento teleosemántico “*top-down*” frente a otro “*bottom-up*”.⁹⁸

La estructura del capítulo es la siguiente. En la segunda parte revisaré la estructura de la objeción a partir del celebre mecanismo de detección de la rana *Rana pipiens*, donde la premisa central es que los contextos de explicación selectiva y, por tanto, de atribución funcional, son transparentes a la sustitución de equivalentes nomológicamente necesarios. En las secciones siguientes cuestionaré las premisas de éste argumento y defenderé que los contextos selectivos son opacos pero que dicha opacidad no es suficiente para elegir entre descripciones divergentes de la función relevante. Mi conclusión positiva será que es una consecuencia del enfoque evolucionista y de la propuesta teleosemántica de Millikan reconocer que hay *grados* de determinación a lo largo y ancho de los linajes, y, por tanto, que hay grados de indeterminación en los sistemas cognitivos representacionales de los organismos, donde el grado de determinación es relativo a las exigencias históricas y ecológicas particulares a cada linaje. Este argumento depende crucialmente de distinguir entre la exigencia de *determinación* de contenido y la exigencia de determinación *unívoca* del contenido, donde la determinación unívoca es característica del contenido *proposicional*. En este sentido, la característica (hiper) intensionalidad de los contextos de atribución de actitudes proposicionales nos ofrece un modelo defectuoso para dar cuenta de sistemas representacionales cuyo contenido no es proposicional.

2. La objeción.

El problema ha sido expresado con el ejemplo del sistema visual la rana *Rana pipiens* (Lettvin et al, 1959).⁹⁹ Confrontada con ciertos estímulos en el entorno, la rana produce un estado perceptual

⁹⁸ En un sentido, una versión teleosemántica “*top-down*” es una que comienza por aplicar la teoría al contexto de la psicología intencional humana (Papineau, 1984, 1993). En oposición, una versión “*bottom-up*” comienza por aplicar la teoría a organismos cognitivamente más simples y va ascendiendo en grados de complejidad (Millikan, 1984, 1993). En otro sentido, una versión teleosemántica “*top-down*” refiere a la idea de que el beneficiario o consumidor relevante en la determinación de la función / contenido de una representación es organismo completo (Millikan, 1984, 1991), en oposición a una versión “*bottom-up*” donde el beneficiario o consumidor es el mecanismo próximo (Neander, 1995). Aquí me refiero a este segundo sentido.

⁹⁹ Originalmente el problema fue inspirado por el célebre artículo de neurofisiología “*What the frog’s eye tells the frog’s brain*” de Lettvin, J. Y., Maturana, H. R., McCulloch, W. S. & Pitts, W. H. que apareció en “*Proc. Inst. Radio Engr*”. 1959, vol. 47 pp. 1940-1951, reimpresso en “*The Mind: Biological Approaches to its Functions*” (Editores: William C. Corning, Martín Balaban, 1968, pp. 233-258). Desde entonces mucha tinta ha corrido sobre el

–una representación interna N –, que causa la captura con su lengua y la ingesta consecuente de lo que captura. En su hábitat natural, la conducta de captura es causada por la detección e ingesta de moscas. Sin embargo, bajo condiciones diferentes, cualquier objeto pequeño, móvil y oscuro puede ser ingerido ya que la elección de comida está determinada sólo por el tamaño y el movimiento del objeto detectado (1959, p, 231).¹⁰⁰ La pregunta es, ¿cuál es la función propia / contenido de N ?

Ciertamente parece haber un recuento adaptativo o etiológico sobre la selección de dichos mecanismos en los ancestros de la rana capaz de arrojar una respuesta determinada a la pregunta arriba planteada. Sin embargo, al tratar de describir la función de N nos enfrentamos a una proliferación de asignaciones funcionales que aparentemente son excluyentes entre sí pero compatibles con el recuento etiológico. Por ejemplo, la función del mecanismo que subyace a N podría ser la de detectar y capturar *moscas* (Sterelny, 1990, Millikan, 1991, 1993), o simplemente detectar *comida* (Shapiro, 1992) o incluso la de detectar *objetos pequeños y oscuros en movimiento* (Neander, 1995), o finalmente, la disyunción *moscas-o-objetos-pequeños-y-oscuros-en-movimiento* (Fodor, 1990). La dificultad estriba en que cada una de estas descripciones alternativas está basada en la teoría etiológica: las ranas ancestrales incrementaron su aptitud por detectar moscas, pero también por detectar comida o también por

tema en filosofía sobre qué determina el contenido de un mecanismo de detección: Dretske (1986); Dennett (1988, 1995); Fodor (1990, 1996); Millikan (1991, 1993); Shapiro (1992); Neander (1995); Goode y Griffiths (1995); Rowlands (1997); Price (1998); Perlman (2002); Enc (2002); Walsh (2002), entre otros.

¹⁰⁰ Otro problema relacionado sobre la indeterminación del contenido concierne al célebre ejemplo de Dretske (1986, 1988). El protagonista es una especie de bacteria en el hemisferio norte que es capaz de orientarse en dirección opuesta a la superficie del agua rica en oxígeno que le resulta tóxica. Para lograrlo, la bacteria cuenta con “*magnetosomes*”, que son pequeños magnetos internos que la empujan hacia el Polo Norte y por lo tanto hacia abajo, lejos de la superficie letal para la bacteria. Bajo este esquema, la función de los magnetos es la de efectuar que la bacteria se mueva hacia el agua libre de oxígeno. Correlativamente, una función del sistema magnético es producir estados o signos que concuerdan con la localización del agua libre de oxígeno. Dadas estas funciones, ¿qué es lo que el empuje de los magnetos representa? Por cuestiones de espacio no podemos abordar este caso pero su solución es similar al del caso de la rana.

Sin embargo, cabe señalar por qué el ejemplo de la rana es más pertinente que el de la bacteria. Así como es oportuno distinguir entre capacidades representacionales humanas y no humanas y entre tipos de contenido intencional (conceptual o no-conceptual, proposicional, etc.) es importante también delimitar qué capacidades cuentan como genuinamente *representacionales* en opción a capacidades de mera *transducción*. La distinción no es clara y distinta a la fecha. Podemos decir que lo que marca la diferencia entre un organismo que posee sólo capacidades de transducción de otro que posee genuinas capacidades de representación es, en parte, la posesión de mecanismos internos sensibles a estados de distintos mecanismos que están dirigidos a aspectos del entorno. La presencia de tales mecanismos internos parece ser necesaria para el tipo de múltiple desacoplabilidad de las representaciones mentales característica de sistemas representacionales cabales y completos como el que subyace a la psicología intencional humana. Ver Cummins et al (2006) para algunas ideas entorno a esta diferencia.

detectar objetos pequeños y oscuros en movimiento. Así, la teoría etiológica no ofrece un criterio para elegir entre estas descripciones alternativas de la función.

La objeción contra la teleosemántica es entonces que individuar a las representaciones mentales a partir de sus funciones evolutivas tiene la consecuencia indeseable de que el contenido de una representación mental queda indeterminado en virtud de que las atribuciones funcionales están ellas mismas indeterminadas. Así, no podemos identificar el contenido de las representaciones mentales con sus funciones propias biológicas porque los contextos selectivos de atribución funcional no exhiben la familiar opacidad que exhiben los contextos de atribución intencional, paradigmáticamente, las actitudes proposicionales (Dretske, 1986, Fodor, 1990, 1998).¹⁰¹ En este sentido, la indeterminación que aqueja a la teleosemántica proviene de la indeterminación que aqueja al teleofuncionalismo: si la individuación de las funciones biológicas no basta para individuar los contenidos intencionales, es porque la individuación teleofuncional no es tan finamente discriminada como lo es la individuación de contenidos intencionales.^{102- 103}

Tenemos entonces que si *N* es una representación particular derivada de un mecanismo representacional *M*, y si es posible ofrecer descripciones intensionalmente¹⁰⁴ divergentes de la

¹⁰¹ Si esta objeción es exitosa, entonces la teleosemántica no logra ofrecer una naturalización completa del contenido semántico, en particular del contenido semántico proposicional característico de estados cognitivos como son las creencias y deseos. Pero esto no sería una objeción para una teoría teleosemántica pura en virtud de que el propósito de dicha teoría es dar cuenta de la normatividad y no de la intensionalidad de la intencionalidad. En especial si consideramos que la putativa intensionalidad no caracteriza a todos los mecanismos cognitivos representacionales, sino sólo aquellos cuyo contenido representacional es proposicional. En este sentido, la intensionalidad caracteriza los contextos de *atribución* de actitudes proposicionales, y por lo tanto no puede ser tomada como una condición necesaria para considerar a una representación *intencional*. Por otra parte, si la objeción resiste el escrutinio crítico, el resultado sería una seria objeción contra la teleosemántica inmodesta y no contra una versión modesta dirigida a dar cuenta de las capacidades representacionales de organismo menos sofisticados que los humanos, por ejemplo, los mecanismo de detección.

¹⁰² Esto permite mostrar la estructura del argumento general que hemos explicitado en la introducción: la viabilidad de la teleosemántica tanto modesta como inmodesta depende de la viabilidad del teleofuncionalismo en biología evolucionista. La idea es que el teleofuncionalismo es al menos una condición necesaria –si bien no suficiente–, para la teleosemántica en el caso de la psicología intencional de actitudes. Pero quizá el teleofuncionalismo sea una condición necesaria y suficiente para la teleosemántica modesta.

¹⁰³ Es de llamar la atención que la putativa indeterminación ha sido formulada en relación a organismos no-humanos cuyos sistemas representacionales son mecanismos de *detección*. Estos estados representacionales no involucran la sofisticación de los mecanismos que presumiblemente subyacen a la capacidad de tener genuinas actitudes proposicionales. Veremos que la exigencia de *determinación* robusta en el sentido de determinación *unívoca* del contenido representacional de una actitud proposicional, simplemente no se aplica a mecanismos de “detección”.

¹⁰⁴ “Descripción intensional” hace referencia al hecho de que ‘mosca’, ‘alimento’, ‘objetos negros y pequeños en movimiento’ tienen la misma *extensión*, a saber, refieren al mismo objeto detectado-capturado, pero difieren en la manera en que el objeto es representado. En otras palabras, lo *qué* es representado es lo mismo en cada caso; la

función de N y de la función de M , entonces la función de M y de N están indeterminadas.¹⁰⁵ El problema se resolvería ofreciendo algún hecho o condición que permitiese elegir alguna descripción como la correcta, esto es, como determinando específicamente el contenido de la representación. Sin embargo, según los críticos del programa teleosemántico, ese “hecho” o “condición” no se encuentra en el repertorio de la teoría etiológica, en particular, en las propiedades selectivas con las que la teoría identifica las propiedades intencionales y semánticas.

¿Cuál es el problema con el recuento etiológico que genera la proliferación de descripciones de la función de N ? Para Fodor (1990, 1998) el principal problema es que la noción relevante de función es histórica y las explicaciones históricas de tipo selectivo, según él, no ofrecen una base para distinguir entre descripciones intensionalmente alternativas pero explicativamente adecuadas de la función de un rasgo. En otras palabras, la indeterminación de la función-contenido es una consecuencia de la transparencia extensional de los contextos selectivos. Esta indeterminación se expresa en el supuesto carácter estrictamente *extensional* de los enunciados que asignan funciones propias (Fodor, 1990, p.71-73 y Fodor, 1998, p. 176). El enunciado ‘la función del mecanismo evolucionado M es la de representar F ’ puede ser sustituido *salva veritate* por el enunciado ‘la función del mecanismo evolucionado M es la de representar G ’ siempre y cuando entre F y G haya una conexión causal necesaria, esto es, soportada por contrafácticos de tal suerte que el enunciado ‘ F si y sólo si G ’ es confiablemente (y nomológicamente) verdadero.

En resumen, la objeción afirma que: a) los contextos selectivos son transparentes a la sustitución de rasgos instanciados por los mismos organismos. b) Esta transparencia implica la indeterminación de las atribuciones funcionales. Y, c), puesto que las atribuciones de estados

diferencia radica en *cómo* es representado. Así, las diferentes descripciones son coextensionales pero difieren en el modo de presentar la extensión, por lo tanto, difieren en el contenido, pues el contenido incluye no sólo una extensión sino un modo de presentar la extensión: una condición que el objeto de la representación debe satisfacer para caer dentro del contenido de la representación. Siguiendo a Carnap (1947) podemos llamar a esta condición la “intensión” de la representación. De ahí que no baste que dos representaciones tengan la misma extensión para tener el mismo contenido, es necesario que tengan además la misma intención. Puesto que el tema que nos ocupa es el contenido *intensional* de la representación, y en principio diferentes intenciones pueden corresponder con la misma extensión, el problema es que la coextensionalidad de las diferentes descripciones no es suficiente para elegir entre ellas a una que fijara el contenido.

¹⁰⁵ Siguiendo el esquema de Millikan (1984), la función de N es una función propia *derivada* (adaptada) de la función propia *directa* (relacional) de M . Puesto que M es un mecanismo cognitivo y N un estado cognitivo, la función propia directa y derivada son funciones propias *relacionales*, esto es, mantienen una “relación estructural abstracta” con un aspecto del entorno.

intencionales no son transparentes, al identificar las atribuciones intencionales con las atribuciones funcionales tenemos que la indeterminación funcional implica la indeterminación intencional. Para responder a esto, a continuación veremos si la premisa ‘a)’ está justificada y si la propuesta de entender los contextos selectivo como contextos *confiablemente* extensionales es pertinente dada la naturaleza histórica de las explicaciones evolutivas.

3. La opacidad causal de los contextos selectivos.

La premisa que afirma que la selección natural es insensible a diferencias intensionales es prematura. Sober (1984) presenta la noción de “selección por” como una noción causal intensional: la selección *por* propiedades no resiste, *salva veritate*, la sustitución de predicados coextensionales.¹⁰⁶ La motivación para introducir las nociones de ‘selección por’ y ‘selección de’ es contrastar los conceptos de *aptitud* y de *selección natural*, en particular, la insuficiencia del concepto de “aptitud global” de un organismo para reflejar la composicionalidad causal de un proceso selectivo (Sober, 1984). La insuficiencia emerge cuando queremos determinar la aptitud de un rasgo en términos de la aptitud promedio del organismo con ese rasgo. La consecuencia indeseable es que:

- (i) dos rasgos instanciados por los mismos organismos tienen el mismo valor de aptitud. Sin embargo,
- (ii) dos rasgos equivalentes en valor de aptitud pueden diferir en su importancia selectiva. Por lo tanto,
- (iii) la mera *equivalencia extensional* de dos rasgos presentes en los mismos organismos a la luz de la aptitud promedio de los organismos con ese rasgo es insensible a la diferencia entre, por ejemplo, que un rasgo sea selectivamente neutral y que otro sea el objeto [“*target*”] de la selección.

¹⁰⁶ Así, de la misma manera en que en el contexto psicológico, alguien puede creer que *P* sin creer que *Q*, aunque *P* si y sólo si *Q*, en el contexto biológico una criatura puede ser seleccionada por ser *X* y no por ser *Y* aunque todos y sólo los *X* sean *Y* (y por tanto haya habido selección *de Y* pero *por* o “a causa” de *X*). Sin embargo, como veremos más adelante, la opacidad causal de los contextos selectivos no es tan fuerte como la opacidad de los contextos de atribución intencional.

En pocas palabras, la conclusión de Sober es que el concepto de aptitud global es causalmente “inerte” puesto que es una *consecuencia* del proceso de selección y no una *causa*.¹⁰⁷ En cambio, el concepto de selección por y contra *propiedades* permite identificar las diferencias en la significatividad selectiva de los rasgos que se aplican a los mismos individuos.¹⁰⁸ Como lo muestra el ejemplo de la “máquina de selección”, la selección *de* propiedades se obtiene como *efecto* del proceso selectivo, mientras que la selección *por* propiedades describe las *causas* de dicho proceso: la propiedad relevante causa el éxito en la supervivencia y reproducción, pero un objeto meramente aumenta su representación como resultado de la selección. La selección natural resulta en la reproducción diferencial de rangos completos de propiedades. Algunas de estas propiedades son meros efectos secundarios del proceso de selección mientras que otras constituyen las causas u objetos de la selección. ¿Qué propiedades son meros efectos del proceso selectivo y cuales son causas de éste? Esto es algo que se reduce a la verdad de ciertos contrafácticos.¹⁰⁹ Por ejemplo (Goode y Griffiths, 1995), el grueso abrigo del oso polar es seleccionado por sus propiedades térmicas, pero puesto que un abrigo caliente, debe ser también un abrigo pesado, también hay selección de la propiedad de ser pesado. Podemos discriminar la propiedad que causa la selección (temperatura) de la que es un mero efecto (ser pesado) contrafácticamente: si el abrigo del oso polar hubiera sido caliente pero no pesado aun así hubiera sido seleccionado. Pero si el abrigo hubiera sido pesado pero no caliente no hubiera sido seleccionado.

¹⁰⁷ Puesto que decir que un organismo individual es más apto que otro no es aún decir *por qué* es más apto, la diferencia conceptual entre la aptitud y la selección natural se mantendría aun si la aptitud causara el éxito en la supervivencia y reproducción (Sober, 1984, p. 102). La oración ‘C1 causó E’ da un imagen más precisa de los hechos causales que la oración disyuntiva ‘C1 o C2 o C3 o...o Cn causó E’. La primera oración es más descriptiva puesto que dice que hay una característica o propiedad por la que hubo selección contra otra propiedad. Tal y como lo presenta Sober, cabe resaltar que esta falta de especificidad no hace al concepto de aptitud un concepto defectuoso, ya que es perfectamente pertinente al nivel de generalidad donde es usado para generar patrones de explicación.

¹⁰⁸ “Pensemos en una “máquina de selección” (Sober, 1984): un contenedor que posee distintos niveles con agujeros cada vez más pequeños conforme van descendiendo los niveles. En la parte superior hay pelotas de dos tipos, unas son amarillas y pequeñas, otras son rojas y de mayor tamaño. Si las dejamos caer a través de los distintos niveles de agujeros, sólo las pelotas pequeñas y amarillas llegarán hasta el nivel más bajo mientras que las pelotas rojas y grandes se quedarán en los niveles superiores. Lo que esto muestra es que hay selección *de* pelotas amarillas pero *por* ser pequeñas. Así, es la propiedad de ser pequeñas y no la propiedad de ser amarillas, esto es, el tamaño más no el color, lo que explica –causa, que esas pelotas hayan sido seleccionadas.

¹⁰⁹ El contraste central para entender la estructura causal de los procesos evolutivos es entre la noción de ‘aptitud global’ y la de ‘selección por’, y radica en que la primera describe la vulnerabilidad de un organismo a la *posibilidad* de un proceso de selección, mientras que la selección por propiedades describe un proceso de selección *actual* (Sober, 1984).

Ahora bien, lo relevante para el problema de la indeterminación es, *contra* Fodor, que el concepto de “selección por” es causalmente *oblicuo*, no resiste la substitución de de predicados coextensivos *salve veritate*: un organismo puede ser seleccionado por tener el rasgo *F* y no por tener *G* incluso si *F* y *G* son extensionalmente equivalentes (son instanciados en el mismo organismo), esto es, tienen el mismo valor de aptitud promedio en la aptitud global del organismo. En resumen, la intuición detrás de la intensionalidad de los contextos selectivos es que la siguiente inferencia (I) no es válida:

- (P1) ‘Los *x* fueron seleccionados por tener la propiedad *F*’
- (P2) ‘*F* si y sólo si *G*’
- (C) ‘Los *x* fueron seleccionados por tener la propiedad *G*’

La invalidez resulta de la verdad de los siguientes enunciados contrafácticos:

- ‘Los *x* serían seleccionados en los mundos posibles donde poseen la propiedad *F* pero no la propiedad *G*’, pero,
- ‘Los *x* no serían seleccionados en los mundos posibles donde poseen la propiedad *G* pero no la propiedad *F*’

Sin embargo, la principal objeción contra la supuesta intensionalidad de la noción de selección por (Fodor, 1990, p.71-73 y Fodor, 1998, p. 176) es que si bien la mera equivalencia extensional entre ‘*F*’ y ‘*G*’ no permite la substitución de ‘*F*’ por ‘*G*’, la equivalencia *necesaria* si lo hace. Así, la correlación entre *F* y *G* no es accidental sino *nomológicamente* necesaria. Podemos entonces validar la inferencia (I) cambiando (P2) por (P3):

- (P1) ‘Los *x* fueron seleccionados por tener la propiedad *F*’
- (P3) ‘*F* si y sólo si *G*’ es una ley (*F* y *G* están nómicamente correlacionadas)
- (C) ‘Los *x* fueron seleccionados por tener la propiedad *G*’

Así, los enunciados que expresan la atribución de una función propia sí pueden ser substituidos por enunciados co-extensionales sin alterar el valor de verdad del enunciado si son confiablemente co-extensionales, esto es, si hay una conexión nomológicamente necesaria entre ellos. Si aceptamos que en el ambiente de las ranas hay una correlación no sólo accidental, sino

confiable entre capturar moscas y capturar pequeños-puntos-negros-en-movimiento el recuento etiológico no es suficiente para determinar cuál es la función tanto de la representación interna como de la acción correspondiente de captura. En este sentido, los contextos de explicación selectiva y, por tanto, de atribución funcional son transparentes a la sustitución de equivalentes nomológicamente necesarios.¹¹⁰

Sin embargo, debemos preguntarnos cuál es la pertinencia, desde el punto de vista de una explicación evolucionista-selectiva, de apelar a escenarios contrafácticos para garantizar la transparencia de los contextos selectivos. Como el propio Fodor (1990, 1998) lo reconoce, apelar a contrafácticos “nomológicamente confiables” es irrelevante dado el carácter *histórico* del proceso selectivo: la selección actúa sobre variaciones actuales y no “posibles” de los rasgos que selecciona. El que estas explicaciones sean irreduciblemente históricas hace patente el carácter constitutivamente contingente de los escenarios y los factores que intervienen en la selección y que no pueden ser soslayadas apelando a “conexiones necesarias” incluso “nomológicamente necesarias”. En otras palabras, la explicación causal de aquello *por* lo que un sistema o estructura es seleccionado supone mencionar el rol causal de ciertas propiedades frente a otras *independientemente* de que tales propiedades estén nomológicamente conectadas.¹¹¹ En este sentido, tampoco está claro cómo determinar el valor de verdad de estos contrafácticos, ya que por un lado, si x no tuviera la propiedad F podría ser seleccionado por la propiedad G , H , S , o T .¹¹² Pero decir que si un rasgo no hubiera sido seleccionado por producir F hubiera sido seleccionado por producir G , H , S o T supone que el rasgo hubiera sido seleccionado de cualquier manera por una razón u otra, lo cual implica una adaptacionismo exagerado.¹¹³ Sólo las

¹¹⁰ Y esto no sólo para un rasgo intencional, como en el caso del mecanismo de detección donde atendemos al fenotipo cognitivo del organismo, sino que esta transparencia se cumple para toda atribución funcional.

¹¹¹ Para Millikan (1991), a diferencia de Fodor (1991), la determinación del contenido es histórica, no contrafáctica. Para Neander (1995), es posible reconciliar ambas cosas: si la historia es algo más que una secuencia de sucesos o una serie de correlaciones estadísticas, esto es, una serie de relaciones causales entre sucesos y las relaciones causales de este tipo dependen de ciertos contrafácticos, podemos concluir que la historia depende de contrafácticos. Sin embargo, no son los contrafácticos *qua* contrafácticos el problema aquí, pues de hecho, la distinción soberiana se explica apelando a escenarios contrafácticos. En cambio, el problema con la propuesta de Fodor es que la historia es irrelevante para contrafácticos nomológicamente necesarios. Y aquí sí hay una divergencia entre explicaciones históricas e explicaciones por leyes. Las relaciones causales de las ciencias históricas, a diferencia de las ciencias nomotéticas, es que no son simétricas. No abundaré más al respecto.

¹¹² En este sentido, las explicaciones teleológicas o etiológicas ofrecen condiciones suficientes más no necesarias para la selección (Fodor, 1998, p. 179).

¹¹³ Esto pasa por alto el hecho de que frecuentemente el costo involucrado en desarrollar mecanismos más sofisticados rebasa el beneficio de tener tales mecanismos (Dawkins, 1982).

variaciones actuales son relevantes y no todas tienen que tener un efecto positivo en la aptitud. En este sentido, los contrafácticos en cuestión no sólo están indeterminados en su valor de verdad sino que tienen altas probabilidades de ser falsos.

Podemos concluir que la noción de que un rasgo puede incrementar la aptitud de un animal en virtud de ciertas propiedades y no de otras sólo tiene sentido en el contexto de variación selectiva actual y no de una lógica o nomológicamente posible. Sólo así es que hay rasgos determinados que pueden ser seleccionados por ciertas propiedades que incrementan más que otras variaciones la aptitud de sus portadores. Así, las teorías teleosemánticas son teorías *actualistas* de la representación (Fodor, 1990, Rowlands, 1997). Esto quiere decir que lo que le da a un estado cognitivo su contenido representacional es la historia *actual* evolutiva de los mecanismos que la producen y consumen y no escenarios históricamente *posibles*.

Hemos visto que la premisa de la objeción que interpreta los contextos selectivos como contextos extensionales, incluso como contextos confiablemente extensionales, no está justificada. Los contextos selectivos (“selección por”) son después de todo causalmente opacos. Ahora veremos si la desactivación de esta premisa es suficiente para desactivar la objeción.

4. La jerarquía niveles de atribución funcional y la falacia del argumento del objeto de selección.

Algunos filósofos como Sterelny (1990), Millikan (1991) y Shapiro (1992) han recurrido a la distinción soberiana para determinar la función de N . Otros, como Neander (1995) y Price (1998) han tratado de resolver esta indeterminación implementado ciertos constreñimientos a las atribuciones funcionales.¹¹⁴ Me concentraré en la primera estrategia, ya que la segunda supone una posición revisionista de la práctica biológica (Walsh, 2002).

¹¹⁴ Price (1998) pretende enmendar la formulación original de la teoría etiológica de Millikan agregando dos condiciones. La condición de “independencia” y la condición de “abstracción”. La primera condición afirma que la función determinada de un rasgo x es el efecto que x tiene *independientemente* de los varios mecanismos de consumo que dependan de él. La segunda condición supone que la función de un rasgo debe ser especificada de manera que se abstraiga de a) la estructura específica del rasgo en cuestión y b) la estructura del mecanismo de consumo que depende de él. Dadas estas condiciones y tomando en cuenta el énfasis de Millikan en la parte que consume, Price concluye lo mismo que Millikan, a saber que el contenido de N es ‘mosca’. Para una crítica de esta respuesta ver Walsh (2002) y Enc (2002).

En los términos del argumento soberiano, la propiedad de ser ‘mosca’ y la de ser un ‘objeto oscuro y pequeño en movimiento’ eran co-extensionales ya que eran instanciadas por los mismos objetos en el entorno donde los ancestros de la rana y los mecanismos de detección evolucionaron. Sin embargo, como enfatiza Millikan (1991, 1993) la propiedad causalmente eficiente en relación a la selección era la de ser ‘mosca’ porque las moscas eran el alimento de las ranas, mientras que la propiedad de ser ‘objetos pequeños...’ no tiene, por sí mismo, ninguna consecuencia sobre la aptitud de las ranas, es un efecto de la selección. Esto justifica para Millikan individuar el contenido de la representación de la rana como ‘mosca’ y no como ‘pequeños puntos negros en movimiento’. En el mismo sentido, Sterelny (1990) presenta la distinción soberiana en términos de enunciados sobre lo que pasaría en posibles escenarios contrafácticos futuros: si los objetos que buscan las ranas fueran ‘objetos pequeños y oscuros en movimiento’ pero no ‘moscas’, no habría selección para detectarlos y atraparlos. Pero si tales objetos fueran ‘moscas’ pero no ‘pequeños objetos oscuros en movimiento’ habría selección para detectarlos y capturarlos. Por lo tanto, el contenido es ‘moscas’. La idea en ambos casos es que muchos objetos que no son moscas cumplen la condición de pequeños oscuros y en movimiento y sin embargo son selectivamente neutrales. La condición biológica en cambio discrimina entre los objetos que tienen valor selectivo y los que no. Moscas, y otras cosas tengan un valor nutrimental o de peligro cumplen con la condición biológica. En ese sentido, la distinción soberiana también puede ser empleada para determinar la función de N como ‘comida’ en lugar de ‘moscas’. Shapiro (1992) afirma que el estado representacional N de la rana tiene el contenido ‘comida’ porque el sub-sistema neuronal del que forma parte es un sistema seleccionado para detectar ‘comida’ en oposición a ‘pequeños objetos oscuros en movimiento’ en virtud de que no habría presión selectiva para estos últimos. Así, es cierto que si los ‘objetos oscuros y pequeños en movimiento’ no fueran ‘moscas’, no serían seleccionados porque las moscas constituyen alimento mientras que los objetos pequeños *qua* objetos pequeños no. Sin embargo, si las moscas no constituyeran alimento, N no hubiera sido seleccionada por representar o detectar moscas. Debemos concluir que ‘comida’ y no ‘moscas’ es lo que causó la selección de N porque es el hecho de que las moscas son alimento para las ranas es lo que explica la presencia y persistencia de los mecanismos que subyacen a N .

Sin embargo, en esta línea podríamos reformular la propuesta de Shapiro y decir que la función de *N* no es ‘alimento’ sino ‘incrementar la aptitud’ ya que si el alimento no incrementará la aptitud no hubiera sido objeto de selección.¹¹⁵ Shapiro podría contestar que aunque los mundos donde el alimento no incrementa la aptitud son lógicamente posibles, son, sin embargo, demasiado lejanos al mundo actual como para constituir una alternativa relevante. Otra respuesta posible es que ‘incrementar la aptitud’ es una descripción demasiado general de *N* como para determinar la función específica de *N*. De hecho, todas las funciones incrementan (o incrementaron) la aptitud por definición (ver (df) p. 32). Si bien esto es cierto, éste problema de la generalidad de la descripción es aplicable a la propuesta de Shapiro, ya que muchos rasgos funcionales en la gran mayoría de especies tienen la función de representar alimento. Si bien la propuesta de Shapiro ofrece un contenido determinado, no ofrece un contenido *suficientemente* determinado como para distinguir a *N* de otras representaciones que tienen, por ejemplo, la función de representar peligro o pareja. Además, es una cuestión empírica si sólo las moscas y ningún otro objeto pequeño, oscuro y en movimiento, tenía valor nutrimental para las ranas.

Tenemos que concluir que, si bien podemos aceptar la opacidad causal de los contextos selectivos, la distinción soberiana no ofrece una solución al problema de la indeterminación aun apelando a la parte consumidora. La razón es que siempre es posible reformular la cuestión en términos de la noción de selección por.¹¹⁶ ¿Es ésta una consecuencia letal para la aplicación del

¹¹⁵ Millikan (1991, 1993) ha dicho que el problema de la indeterminación se soluciona apelando no sólo a la distinción soberiana sino a la distinción entre los mecanismos que producen *N* y los que consumen *N*, siendo esto últimos los que determinan el contenido de *N*. La idea es que el contenido depende de cómo el mecanismo consumidor interpreta (usa) la representación. La determinación depende de la salida conductual y no de la entrada informacional. El contenido es la condición bajo la cual la conducta resultante sería apropiada, independientemente de si las condiciones actuales que causan la representación son de ese tipo. En el caso de la rana esto quiere decir que la parte que consume la representación ‘mosca’ es el mecanismo de captura como su sistema digestivo porque ‘moscas’ eran lo que las alimentaba en el entorno donde evolucionaron las ranas ancestrales y por tanto, las moscas eran la presión selectiva que fijó la función de los mecanismos responsables de capturarlas. Ahora bien, el problema con apelar al “consumidor” es que es posible preguntar qué determina la función propia del mecanismo consumidor, y si la respuesta depende de la distinción soberiana, entonces en principio la función propia del consumidor también está indeterminada pues es posible reformular la función del consumidor de *N* como de *N* misma. En otras palabras, apelar al consumidor, más que solucionar el problema, lo transfiere. Por otra parte, también es posible que otros mecanismos en la rana consuman la representación; más aún, que haya una serie de consumidores más o menos próximos de la representación lo cual generaría distintos contenidos determinados para la misma representación producto de pluralidad de consumidores que en principio dicha representación podría tener.

¹¹⁶ En cuanto a la posibilidad de que el contenido de la representación de la rana sea disyuntivo (Fodor, 1990), tanto Millikan (1991) como Sterelny (1990) y Griffiths y Goode (1994) rechazan que este sea un problema genuino para la propuesta teleosemántica. Principalmente porque las formulaciones disyuntivas no juegan ninguna papel en las explicaciones evolutivas y selectivas, sino en cambio, son empleadas en contextos epistémicos como premisas en

modelo etiológico a una teoría representacional de la cognición?, ¿cuál es el problema con emplear el argumento del objeto de la selección para elegir una descripción como la correcta?

Para Goode y Griffiths (1995) que en este uso de la distinción soberiana, descripciones complementarias de la función de un rasgo son erróneamente presentadas como excluyentes. Lo que este uso de la distinción ignora es que los procesos de selección pueden ser descritos a distintos niveles teóricos.¹¹⁷ Así, la proliferación de funciones que da pie al “problema de la indeterminación” surge de la existencia de múltiples niveles teóricos en el estudio de los organismos y su evolución.¹¹⁸ El uso falaz de la distinción es entonces entre propiedades selectivas a distintos niveles. En cambio, la distinción es relevante *sólo* entre propiedades selectivas o descripciones alternativas de estas propiedades selectivas al mismo nivel teórico. Dicho esto, podemos organizar las distintas descripciones de la función propia de *N* en una jerarquía de niveles que corresponden a distintos niveles teóricos en la biología de la evolución, cada uno con una manera propietaria de describir *N* (Neander, 1995, Goode y Griffiths, 1995):

N contribuye a la aptitud¹¹⁹

Para mediante

N ocurre en respuesta a comida¹²⁰

inferencias deductivas que no involucran mecanismos y por lo tanto resultan irrelevantes desde el punto de vista causal (Millikan, 1991): capturar ‘moscas-o-abejas’ o ‘moscas o pequeños-objetos-negros-móviles-etc’ (Fodor, 1990) no es algo que pueda contribuir a la proliferación del linaje de las ranas.

¹¹⁷ La idea es que emplear la distinción entre distintos niveles de explicación representa equivocadamente descripciones evolutivas complementarias como si estuvieran en competencia o como si sólo un nivel fuera relevante desde el punto de vista de la selección. Dicho esto, evitaré tomar partido en el extenso debate sobre niveles de selección.

¹¹⁸ La clasificación de un mismo dominio de objetos en una serie de vocabularios cada vez más abstractos permite acceder a generalizaciones que no emergen a niveles inferiores en virtud de la múltiple realizabilidad de las propiedades que intervienen en generalizaciones a niveles superiores.

¹¹⁹ La ‘aptitud’ es el nivel teórico más abstracto. En este nivel podemos explicar la evolución del estado *N* mediante generalizaciones del tipo empleadas en genética de poblaciones y teoría de juegos (“nivel poblacional dinámico”). La explicación se realiza mediante una ecuación que representa las frecuencias y la aptitud de *N* y de variaciones de *N*. Esto permite computar la frecuencia de *N* en la población después de un número dado de generaciones. Estos modelos no requieren más que un supuesto de heredabilidad (genética) y valores de aptitud. Siguiendo a Sober (1984) podemos llamar a este nivel de explicación biológica el nivel de los efectos o consecuencias ya que estas explicaciones dependen de leyes sobre las consecuencias de diferencias en la aptitud en lugar de explicar la fuente o las causas de estas diferencias.

¹²⁰ Como señalan Goode y Griffiths (1995), este nivel corresponde con generalizaciones ecológicas del tipo empleadas en explicaciones adaptativas. A diferencia del nivel poblacional, las generalizaciones relevantes explican la fuente de las diferencias de aptitud. Si bien estas explicaciones no son tan abstractas como las que ocurren al nivel poblacional dinámico, puesto que se aplican a cualquier organismo que se enfrenta a un problema adaptativo en

Para mediante
N ocurre en respuesta a moscas ¹²¹

Para mediante
N ocurre en respuesta a pequeños objetos oscuros en movimiento ¹²²

La idea es que lejos de generar un conflicto entre las descripciones, describir el rasgo de la rana a cada nivel permite acceder a distintas generalizaciones complementarias sobre la evolución de *N* (genética de poblaciones, ecología, historia natural, neuroanatomía). De esta manera, la distinción permite establecer qué propiedades intervienen en generalizaciones útiles sobre la selección *a ese nivel*. El problema con el “problema de la indeterminación” es que la distinción se aplica para privilegiar una propiedad sobre otra. La falacia radica en que una explicación a un nivel superior es considerada como una alternativa a una explicación a un nivel inferior en lugar de ser considerado un complemento (incluso en el sentido fuerte en que el nivel inferior realiza el nivel superior de acuerdo a su función). En las propuestas revisadas al principio de este apartado, la falacia se hace patente en el hecho de que en cada caso parece no haber selección o bien a niveles superiores o bien a niveles inferiores.¹²³

tanto ocupa un rol ecológico, aún están muy lejos de la historia natural de una especie particular. Un ejemplo de explicación a este nivel es el de la teoría de la señal-detección (ver Godfrey-Smith, 1991).

¹²¹ Este nivel es consecuencia de que las descripciones ecológicas generales son realizadas en casos particulares por linajes particulares de organismos o rasgos. Hull (1987) llama a esto “actores genealógicos en roles ecológicos”. Podemos decir que el rasgo en la rana que detecta la presencia de moscas es la solución de ese linaje particular al problema de la señal-detección. Este nivel de explicación corresponde a la historia natural, esto es, al estudio de la evolución actual de un sistema particular históricamente delimitado. Así, la afirmación de que la clasificación ‘mecanismo de captura de moscas’ donde ‘moscas’ denota un grupo histórico particular equivale a la afirmación de que hay un patrón suficiente en los sucesos históricos particulares que permiten generalizaciones útiles. La idea es que la historia natural particular añade información relevante al modelo ecológico más amplio y abstracto. La diferencia puede ser entendida como una diferencia entre una explicación causal detallada y una idealización (Goode y Griffiths, 1995, p. 107).

¹²² El último nivel es el neuroanatómico y el de otras disciplinas que describen las capacidades físicas actuales de los rasgos cognitivos. Sin embargo, estas disciplinas apelan a nociones teleológicas en la medida en que distinguen las capacidades físicas de los rasgos que corresponden a sus funciones biológicas propias de aquellas que son meros efectos secundarios. La explicación a este nivel atiende la manera precisa en que la rana está organizada internamente al mostrar como este arreglo u organización resulta en la detección e ingesta de objetos que tienen cierta apariencia.

¹²³ Esto arroja luz sobre la indeterminación del valor de verdad de enunciados contrafácticos sobre sucesos selectivos. Las propiedades seleccionadas a cierto nivel dependen de que otras propiedades son seleccionadas a otros niveles.

5. Niveles teóricos y explicación intencional: el debate por el beneficiario.

Hemos aceptado que existen múltiples niveles legítimos y complementarios de atribución funcional de un mismo rasgo y que esta multiplicidad de niveles responde a distintas exigencias de explicación. La consecuencia relevante para el problema que nos ocupa es que el argumento del objeto de la selección es válido sólo entre propiedades al mismo nivel teórico. Puesto que los niveles son complementarios, la exigencia de elegir *un* nivel como el que determina la función de *N* no está justificada. La exigencia de elegir entre distintas descripciones de la función de *N* está justificada sólo entre descripciones divergentes en el mismo nivel teórico de explicación. La pregunta es, ¿cómo hacer compatible esta multiplicidad con la exigencia de determinación que se sigue de la noción misma de contenido intencional? En otras palabras, ¿cuál es la relación entre la idea de distintos niveles jerárquicos y la idea de fijar nivel pertinente de explicación *intencional*?¹²⁴

Podría ser que esta tensión no tuviera una solución evidente o que ni siquiera necesite una. Quizá sólo expresa eso, una tensión, entre otras, entre la herencia “interpretativista” de la noción de intencionalidad, y su aplicación en el contexto científico. Así, en continuidad con la conclusión pluralista podríamos decir que todos los niveles son relevantes: podemos organizar las distintas descripciones del contenido en series ordenadas que corresponden a series de niveles teóricos relevantes para la psicología cognitiva –neuroanatomía, neurofisiología, neuropsicología, psicología cognitiva, etología y ecología de la conducta, cada uno con una manera propietaria de clasificar el contenido de la representación respondiendo a diferentes exigencias de explicación sin que este pluralismo implique indeterminación: no sólo no hay *una* descripción relevante sino que *cada descripción relativa a cada nivel está determinada a ese nivel*.¹²⁵ Dada esta multiplicidad y esta complementariedad de niveles de atribución funcional, la presión por elegir un nivel, sea el de orden superior o el de orden inferior, se vuelve espuria. El alcance de esta

¹²⁴ Si bien todos los niveles son legítimas fuentes de atribución *funcional*, no todos los niveles tendrían que ser, *ipso facto*, fuentes legítimas de atribución *intencional*. ¿Qué nivel de atribución intencional es el relevante –genética de poblaciones, ecología, historia natural de una especie particular, neuroanatomía?

¹²⁵ Ahora bien, ¿quiere decir esto que el contenido o función de *N* varía según el nivel de explicación o más bien que el contenido es el mismo pero la descripción del contenido es la que varía de acuerdo al nivel de explicación? Pero si hay más de un nivel de determinación, ¿cabría entonces hablar de múltiple-determinación o sobre-determinación? Lo que hay son distintas descripciones –cada una determinada a un nivel, de la misma función. ¿Por qué esta exigencia por elegir una descripción? Una respuesta posible es una confusión entre la exigencia de determinación a secas y la de determinación *unívoca*.

postura depende, sin embargo, de la integración de las neurociencias, la psicología cognitiva y la ecología de la conducta.

En cambio, si decidimos responder a la exigencia de fijar un nivel como el relevante desde el punto de vista de la explicación intencional, necesitamos un argumento que justifique la elección de alguna descripción de *N* que no dependa de la distinción soberiana, sino que apele al nivel que resulte más compatible con la “explicación intencional” de lo que la rana hace. Ahora bien, lo que cuenta como “explicación intencional” depende crucialmente de cómo decidamos individuar el contenido, esto es, o bien de manera amplia (externista) o bien de manera estrecha (internismo). Tal parece que si decidimos honrar el nexo ecológico representación-acción-entorno que parece anidado en la idea misma de explicación intencional, el contenido relevante es amplio.¹²⁶ Esto implica que una psicología intencional científica incluiría generalizaciones al nivel ecológico-histórico. Sin embargo, si decidimos que el nivel relevante es el nivel neurológico (Neander, 1995), no queda claro que honremos el nexo ecológico que intuitivamente supone la idea misma de una explicación intencional. La razón es que a ese nivel, el contenido es descrito sin hacer referencia al objeto en el entorno –moscas, sino a la configuración interna de dicho objeto en organismo –pequeños objetos negros en movimiento.¹²⁷

Una estrategia es ofrecer razones para individuar el contenido en términos de la proximidad o distalidad del estímulo. ¿Cuál es el estímulo relevante para determinar el contenido de *N*? Por una parte, el contenido de *N* podría ser el estímulo próximo, el patrón *en la retina* de la rana que puede ser descrito como ‘puntos negros de cierto tamaño en cierto movimiento’. Por otra parte podría ser el estímulo distante, *algo en el entorno* de la rana, por ejemplo, ‘moscas’. Ahora bien, ¿cómo conecta la cuestión de la proximidad-distalidad con el argumento del objeto de la selección? Puesto que para la teleosemántica el contenido / función de una representación es aquello por lo que los mecanismos representacionales fueron seleccionados, el contenido no puede ser pensado independientemente de un beneficiario. En este sentido, la disputa por el

¹²⁶ En el capítulo anterior señale que la propuesta teleosemántica, al menos la de Millikan, constituía una justificación biológica de un externismo radical. En este apartado defenderé esta postura.

¹²⁷ Con esto no quiero decir que un contenido estrecho sea intrínseco, esto es, que excluya la relación con el entorno. A lo que me refiero es al hecho de que en éste caso diferentes objetos en el entorno pueden instanciar la representación ‘pequeños objetos negros en movimiento’. Así, no sólo las moscas corresponden con ese contenido, y las moscas son el objeto en el entorno que constituye alimento para la rana.

estímulo próximo o distante es una disputa por el *beneficiario* relevante de la representación.¹²⁸ Así, el beneficiario más próximo es el propio mecanismo representacional *M* que subyace al estado representacional particular *N*. Podemos correlacionar al beneficiario próximo con el estímulo próximo, el patrón en la retina. Por otra parte, el beneficiario distante es el organismo completo *O* del que *M* es un rasgo (tipo) igual que *N* (caso). La correlación en este caso es entre el beneficiario último *O* y el estímulo distante: aquello en el entorno que posee un valor real para el organismo. La pregunta es, ¿cuál es el beneficiario relevante de la representación tal que el contenido *intencional* de la representación se fije en relación a su perspectiva? En otras palabras, ¿relativo a qué beneficiario o consumidor debemos describir el contenido de *N* tal que hagamos justicia a la intencionalidad de *N*?, ¿*O* u *M*?

El criterio de Millikan (1984) para elegir entre estas alternativas de lo que *N* se supone que debería representar está planteado en términos indicativos e imperativos, pues para ella *N* es ambas cosas ‘mosca ahí’ y ‘capturar mosca’. (i) En cuanto representación imperativa, lo representado o la condición normal es el *último* miembro de la serie de cosas que se supone que debería representar. (ii) En cuanto representación indicativa, lo representado o la condición normal es aquello que debe ser mencionado al dar la explicación normal más próxima para la ejecución propia *completa* del mecanismo consumidor en tanto adaptado a la representación. De esto se sigue que *N* es una representación de un aspecto *del mundo o del entorno* del animal y no un patrón de estimulación de la retina (‘objetos pequeños en movimiento’). La idea es que al dar la explicación normal más próxima de cómo el mecanismo consumidor de *N* sirve al conjunto completo de sus funciones propias, sólo el hecho de que alguna variable o aspecto en el entorno del animal es representado por *N* debe ser mencionado. Para Millikan la manera de describir la función propia es entonces mediante *la explicación normal más próxima de la ejecución propia completa* (1984, p. 100). Esto quiere decir que la elección de la descripción de la función se hace en relación al beneficiario último, esto es, *O* y no a *M*. Esto implica que la función de la acción de captura no es la de capturar pequeños y oscuros objetos en movimiento sino moscas pues las moscas son el tipo de estímulo externo que explican el beneficio del organismo, por qué los ancestros de *O* proliferaron en virtud de que del que el rasgo *M* producía el efecto *N*.

¹²⁸ En términos de Millikan, la disputa es por el consumidor de la representación, tal que el contenido de ésta esta determinado por el consumidor.

Esta propuesta “*top down*” o del “beneficiario último (organismo)” de Millikan ha sido criticada por Neander (1995)¹²⁹ quien en cambio propone que debemos buscar al nivel más bajo dentro de la jerarquía de niveles teóricos. La motivación para elegir dicho nivel es que, según ella, esto refleja la práctica biológica; el nivel más bajo de *N* antes de pasar a los niveles de sub-*N*. La idea es que los biólogos atienden el nivel más bajo de atribución funcional porque a ese nivel la explicación de las capacidades-disposiciones del sistema pueden ser analizadas en términos no-teleológicos sino en términos de un análisis funcional *a la* Cummins, (1975, 2002).¹³⁰ Así, para ella el estímulo relevante es próximo y el beneficiario es el mecanismo o rasgo *M* en la rana. ¿Qué razones tenemos para elegir *O* u *M* cómo el beneficiario / consumidor relevante?

Mientras que apelar a los niveles muy altos, por ejemplo, incrementar la aptitud, tiene la dificultad de que las diferencias entre las funciones de distintos rasgos se borran -puesto que por definición todas las funciones propias tienen el propósito de incrementar la aptitud-, la dificultad de apelar como Neander al nivel más bajo es que la idea misma de una relación intencional se borra, si bien no en el sentido de perder la normatividad característica, sí en el sentido del “acerca de” en un aspecto *en el entorno*. Si aceptamos que la teleosemántica reivindica parte de las intuiciones externistas de la psicología intencional en términos eco-evo-semánticos, la propuesta de Neander viola una condición mínima para que el mecanismo en cuestión cuente como un caso de representación intencional, a saber, que haya una relación con un aspecto *del* entorno. En este sentido, la propuesta de Neander nos deja con un contenido internista, estrecho, donde no hay mención del objeto distante, la “causa final”, esto es, el efecto complejo que deviene en lo que llamamos “acción intencional”.¹³¹

¹²⁹ La crítica es que si nos concentramos sólo en lo que ha beneficiado actualmente al organismo para determinar el contenido de la representación corremos el riesgo de contra-intuitivamente exigir demasiado de las capacidades representacionales de tales mecanismos (Neander, 1995, Jacob, 1997) en virtud de que tales descripciones del contenido son demasiado abstractas y generales. La cuestión es si podemos realmente atribuir a los organismos en cuestión la capacidad de representarse sólo lo que es “bueno para ellos”.

¹³⁰ La idea es que al nivel más bajo la ventaja es que se conserva la normatividad inherente la teleología justo en la proximidad de los niveles no-teleológicos donde la normatividad desaparece (pero ver Godfrey-Smith, 1993, p. 200). Para Neander (1995) el beneficio es que es posible integrar la explicación teleológica con la explicación mecánica.

¹³¹ En este contexto es relevante tanto la ya mencionada (capítulo segundo) distinción entre funciones propias “directas” y funciones propias “adaptadas derivadas” (Millikan, 1984), como la distinción entre reglas de proyección próximas o algorítmicas y reglas de proyección distantes u organizmicas (Millikan, 1993, p. 219). La regla de nivel inferior o “algorítmica”, que es producto de la función propia derivada-adaptada de la representación, permite

La idea de Millikan es que el nivel relevante para hablar de la *intencionalidad* de la representación es este nivel superior en la medida en que el contenido así definido incluye un aspecto en el entorno y la acción derivada de la representación. En este sentido, la versión teleosemántica de Millikan justifica un contenido externista y como tal la individuación se da al nivel donde las generalizaciones relevantes son ecológicas e históricas sin por supuesto negar que sea realizada al nivel neuronal. Además, la propuesta del beneficiario último de Millikan es completamente consistente con la existencia de distintos niveles de atribución funcional. De hecho, desde esta visión pluralista, lo que el diagrama de Neander, Griffiths y Goode del apartado anterior muestra, es que las funciones propias de niveles más altos se *realizan* mediante la ejecución exitosa de las funciones propias de los niveles o mecanismos inferiores. Los organismos poseen mecanismos con diferentes y múltiples funciones propias que se relacionan entre si en términos de medios-fines. Así, el beneficio relativo al organismo depende de la ejecución propia o normal de la funciones a niveles sub-organísmicos¹³² además de que las condiciones *en el entorno* sean normales. De esta manera, si bien la disfunción puede ocurrir a varios niveles y por distintas causas, lo relevante en el contexto de las explicaciones intencionales es la disfunción de las respuestas complejas que podemos caracterizar como acciones o conductas causadas por la representación y no la disfunción de los mecanismos, a niveles inferiores, que subyacen a estas respuestas complejas. Esto es fiel a la idea de que los organismos poseen rasgos cognitivos que representan aspectos de su entorno para generar respuestas apropiadas que intervienen en el entorno. La propuesta de Millikan reivindica este elemento de las explicaciones intencionales al construir el contenido representacional de un estado cognitivo en relación a la salida conductual y no a la entrada informacional. Así, las condiciones de corrección son amplias porque en el caso *N* el propósito es producir y dirigir la acción de captura.

Sin duda, atribuir una función propia a una representación mental es atribuir una función a nivel neuronal, los mecanismos cognitivos y representacionales son mecanismos neuronales. Pero

explicar la respuesta de la rana al estímulo próximo. La regla de nivel superior u organísmica corresponde a la función propia directa de los mecanismos *M* que lo producen. En estos términos, Millikan da prioridad a las funciones propias relacionales *directas* de los mecanismos relevantes en la *determinación* del contenido de una función propia representacional *M* = mosca mientras que las funciones propias *derivadas-adaptadas* serían las que hacen a la determinación relativamente “unívoca”, *N* = esta-mosca-ahora.

¹³² Rowlands (1997) habla correlativamente de funciones *organísmicas* o de nivel superior y *algorítmicas* o de nivel inferior.

no solamente. El nivel neuronal de atribución es el nivel que realiza la representación mental en el sentido del *vehículo* representacional, pero el *contenido* de la representación no es idéntico a la representación misma de la misma manera que la función del corazón no es idéntica a su estructura.¹³³ En este sentido, el nivel de atribución relevante para el *contenido* de la representación es ecológico y no meramente neuronal. Así, los mecanismos cognitivos representacionales son, en última instancia mecanismo ecológicos. La razón es que la estructura o forma de un rasgo, por sí misma, no explica la función del rasgo pues la relación histórica con el entorno o contexto ecológico es constitutiva de la individuación funcional. En particular si el rasgo es cognitivo, la historia y ecología son indispensables, pues los sistemas cognitivos conectan en formas específicas al organismo con su entorno. La consecuencia central es que, puesto que la individuación de una función propia es resultado de un mecanismo de selección, la individuación de la función propia de un rasgo depende de los detalles del medio ambiente y de las relaciones históricas entre caracteres biológicos y contextos ecológicos, en particular, los entornos selectivos (Brandon, 1990). Si aceptamos que la teleosemántica implica esta versión biológica de un externismo-histórico, el nivel neuroanatómico sólo nos deja con un contenido subdeterminado por el objeto evolutivamente relevante en el entorno (ver nota 31), y por tanto, con un contenido no lo suficientemente amplio, esto es, un contenido capaz de dar cuenta de la dirección a un objeto *en el entorno*, condición que hemos aceptada como constitutiva de un contenido intencional en la medida en que la el contexto de evaluación es la respuesta compleja. En este sentido, el nivel ecológico parece más apropiado pues por definición incluye la relación con el entorno. Tenemos entonces una teoría (psico)semántica naturalista modelada en torno a las disciplinas del evo-eco, lo que podríamos llamar una *ecología cognitiva de la conducta* construida en términos de una teoría representacional de la cognición.¹³⁴

¹³³ Esto se debe a que, en principio, distintas representaciones pueden tener el mismo contenido de la misma manera que distintas estructuras biológicas pueden tener la misma función propia si juegan el mismo rol ecológico en entornos diferentes.

¹³⁴ Ahora bien, es cierto que las ranas responden más frecuentemente a objetos pequeños oscuros en movimiento que a moscas. Pero la razón es que responder ante tales objetos es un medio para poder capturar moscas. Esto sería un problema sólo si la noción de “normalidad” relevante fuera una noción estadística. Así, en condiciones normales –históricamente óptimas, las ranas son capaces de capturar moscas, si bien carecen de la capacidad de discriminar moscas genuinas y meros objetos pequeños y oscuros en movimiento.

6. Conclusión: grados de determinación, “actos de identificación” y pluralismo.

El problema de la indeterminación suele presentarse como una objeción a la teleosemántica *tout court*. Sin embargo, es necesario hacer algunas distinciones. La putativa indeterminación puede afectar a (i) el teleofuncionalismo, (ii) la teleosemántica modesta, o (iii) la teleosemántica inmodesta. Procederé en orden inverso.

En la medida en que el problema de la indeterminación recurre a la característica (hiper)intensionalidad de los contextos donde atribuimos actitudes proposicionales, el problema sí afecta a (iii): la intensionalidad de los contextos selectivos no basta para dar cuenta de la (hiper)intensionalidad de los contextos mentales.¹³⁵ ¿Cómo puede una versión inmodesta de la teleosemántica enfrentar a este problema? La teleosemántica inmodesta puede, *prima facie*, sortear el problema de la indeterminación afirmando que la exigencia de determinación “unívoca” es satisfecha por condiciones relativas al tipo de sistema biológico en cuestión, en este caso, a condiciones específicas de los sistemas cognitivos de orden superior característicos de la psicología intencional (Millikan, 1989, 1991, 1995, Papineau, 1998, Walsh, 2002). La idea es que es natural pensar que, dado que en los sistemas cognitivos que subyacen a la psicología intencional la exigencia de determinación es extremadamente alta –como muestra la característica (hiper)intensionalidad–, estos sistemas han evolucionado en respuesta a estas exigencias y por ende poseen mecanismos con funciones propias para responder a estas exigencias.¹³⁶

Por ejemplo, en la teoría de Millikan (1984), los sistemas cognitivos humanos –a diferencia de sistemas de mera detección–, exigen constreñimientos específicos que responden a la capacidad de manipular racionalmente las representaciones mentales en procesos inferenciales

¹³⁵ La idea es que la intensionalidad de los contextos intencionales es distinta y más aguda que la intensionalidad de los contextos selectivos. La razón es que, mientras lo primeros son contextos mentales y por tanto involucran representaciones, los segundos son contextos histórico-causales y ciertamente no involucran representación (la selección natural es “ciega”). Podemos hablar entonces de *grados* de intensionalidad y correlativamente de grados de determinación.

¹³⁶ Manifestaciones de estas funciones se hacen patentes en fenómenos cognitivos complejos como la presencia de procesos racionales (teóricos y prácticos), la desacoplabilidad de las representaciones mentales con respecto a la producción de acciones y la presencia del lenguaje con el característico aparato referencial que implica y que caracteriza la comunicación e interpretación humana.

lógicos y prácticos. En este sentido, la univocidad en la determinación del contenido representacional responde a las exigencias de los sistemas que consumen la representación. Las creencias humanas, en este sentido, están tan finamente discriminadas puesto involucran lo que ella llama “actos de identificación” (Millikan, 1991, 163) y que no están presentes en organismos no-humanos, esto es, organismo no racionales ni lingüísticos. Por ejemplo, la danza de las abejas no contiene elementos denotativos* puesto que las abejas consumidoras de la danza no identifican los referentes de estas danzas sino que meramente reaccionan de manera apropiada (Millikan, 1984, p. 71). En cambio, los sistemas o mecanismos que subyacen a las creencias y deseos tienen la función propia relativa a ese sistema de *identificar* la referencia de la representación, y es este acto de identificación el que produce la univocidad característica del contenido de estas representaciones mentales.¹³⁷ Así, a diferencia de las abejas, nosotros los humanos *identificamos* el objeto al que las abejas están dirigidas *como* ‘néctar’ usando nuestro esquema lingüístico-conceptual con su característico aparato denotativo-referencial.¹³⁸

Generalizando esta idea, podemos decir que el grado de determinación es relativo al usuario o consumidor de la representación.¹³⁹ Donde el consumidor o usuario es tanto el tipo de organismo que emplea la representación –la especie en cuestión, como los distintos sistemas cognitivos y de control *en* el organismo que hacen uso de la representación en diversos procesos

¹³⁷ Esto implica que la teleosemántica no sostiene y no debería sostener que el contenido representacional del icono intencional de la rana puede ser capturado por el contenido conceptual de una oración en inglés.

¹³⁸ Millikan (1984) presenta una intrincada teoría del “acto de identificación” que aquí sólo mencionaré en líneas muy generales: El acto de identificación del valor real de un elemento de la representación intencional es el acto de saber qué en el mundo dicho elemento mapea o representa (Millikan, 1984, p. 240). Un acto correcto de identificación es tal que para ser debidamente ejecutado requiere al menos de *dos* representaciones intencionales o de dos modos de representación intencional, en donde uno de los elementos de cada una de estas representaciones tiene el mismo valor real que uno de los elementos de la otra representación, tal que un acto correcto de identificación es ejecutado por un mecanismo de consumo que emplea estas representaciones conjuntamente para poder ejecutar su función propia, donde la explicación normal de la ejecución propia de dicha función hace referencia al hecho de que el valor real de ambos elementos es el mismo. En este sentido, dicho mecanismo beneficia al organismo porque dichos elementos representan *la misma* condición en el mundo. Así, en la medida en que el acto de identificación opera sobre pares de representaciones intencionales, dichas representaciones identifican las mismas variantes en el mundo, variantes que se corresponden con variantes en los elementos que componen a ambas representaciones.

¹³⁹ Por su parte, Griffiths y Goode reconocen la multiplicidad de niveles de atribución, la legitimidad de cada nivel y la complementariedad entre niveles, pero no hablan de grados de determinación relativos a la especie en cuestión. Si agregamos esto, no sólo podemos concluir que la indeterminación es genuina pero inofensiva, sino que la supuesta objeción de la indeterminación descansa en extrapolar un grado de determinación, a saber, el relativo a la especie humana, como criterio para todo tipo de contenido representacional.

cognitivos.¹⁴⁰ Así, el requisito de “univocidad” proviene de condiciones relativas a la cognición humana. En particular a la característica opacidad o (hiper)intensionalidad de los enunciados que atribuyen actitudes proposicionales. Consecuentemente, exigir de todo mecanismo representacional determinación unívoca es imponer una medida local, relativa a nuestra especie, a la idea misma de la intencionalidad o de representación cognitiva.

(ii) En esta línea, el caso de la rana es un contexto en el que no parece tener ninguna utilidad evolutiva realizar esa fina discriminación de contenidos, esto es, la identificación *unívoca* de lo representado y por lo tanto, desde la perspectiva de la evolución de su sistema cognitivo, no tiene sentido exigir ese tipo de funciones a la hora de dar el recuento etiológico de dichos sistemas. Fodor (1990, 1998) no considera la posibilidad de que la putativa indeterminación, lejos de ser un problema del recuento teleosemántico modesto, sea por el contrario una condición normal de los mecanismos de detección de la rana. En otras palabras, la premisa implícita e injustificada en el argumento de Fodor es que todos los mecanismos representacionales *qua* representacionales deben exhibir la opacidad característica de los estados intencionales de la psicología humana. Sin embargo, como ya he dicho, esto es una propiedad del tipo de contenido que opera en el contexto cognitivo humano y no una propiedad del contenido *simpliciter*. Si aceptamos hablar de grados de determinación o de grados de intensionalidad como lo permite la perspectiva evolutiva, esta exigencia no está motivada. Lo que la rana necesita representar, dada su historia evolutiva, no requiere la fina discriminación que encontramos en las actitudes proposicionales. El tipo de presiones selectivas que explica la presencia de los mecanismos de detección en la rana no requiere de ese nivel de sofisticación, a saber, la identificación del referente de *N*. En este sentido, la objeción de Fodor implica un antropomorfismo incompatible con la perspectiva evolucionista.¹⁴¹

¹⁴⁰ En esta línea, como afirma Walsh (2002, p. 323) resulta iluminador uno de los elementos centrales del análisis funcional de Cummins, a saber, que una función está determinada sólo *relativamente al sistema que la contiene*. Así como las funciones biológicas son adscritas relativamente al sistema en el que contribuyen, las funciones de los estados intencionales están determinadas relativamente al sistema en que operan, esto es, el sistema de la psicología intencional de creencias y deseos (Papineau, 1993, 1998). En este sentido, si bien encontramos notables dificultades en atribuir una función unívoca a la representacional interna de la rana, una vez que consideramos el rol de los estados intencionales en la economía cognitiva de un agente racional el problema desaparece.

¹⁴¹ El requisito de “univocidad” parece responder más a una presión que hace extensivo el carácter digamos “ultradeterminado” del contenido de creencias y deseos a mecanismos cuyas representaciones intencionales no son de este tipo. La rana genuinamente representa, y en ocasiones se equivoca, sólo no lo hace vía un contenido

(i) Finalmente, es fundamental distinguir entre la *pluralidad* de funciones propias que corresponden a un sistema, mecanismo, entidad o rasgo biológico y la posible *indeterminación* de dichas funciones. La mayoría de los rasgos funcionales poseen un gran número de efectos cada uno de los cuales hacen algún tipo de contribución a la supervivencia y reproducción donde cada uno es, *prima facie*, un candidato posible a ser la función.¹⁴² En este sentido, como dice Millikan (2002), no existe *la* función propia de algo.¹⁴³ En consecuencia podemos afirmar que la teoría etiológica simplemente puede eludir el problema de la indeterminación aceptando tanto la multiplicidad de funciones que puede tener un rasgo, como que la legitimidad de ofrecer distintas formulaciones de la función o funciones de un rasgo. La idea es que es posible hacer formulaciones más o menos próximas o más o menos generales de la función o funciones propias de algo y donde la pertinencia del tipo de formulación está sujeta a consideraciones relativas al nivel de explicación (Millikan, 1991, p. 161). Así, dado el rol explicativo que juegan las asignaciones funcionales en la biología, podemos concluir que no existe el requisito de que tales asignaciones sea completamente determinadas o unívocas: la multiplicidad y el pluralismo resulta endémica a la función evolutiva. Aceptar esta pluralidad no implica aceptar la indeterminación. El corazón tiene la función de bombear sangre, pero también la función de bombear y la función de bombear sangre a través del cuerpo. Estas diferencias son diferencias no en las *funciones* de los órganos sino en las formulaciones con las que *describimos* estas funciones, y como tales, estas diferencias son irrelevantes desde el punto de vista biológico y ciertamente desde el punto de vista del corazón.

proposicional. Por ello hay que tener claro que lo que interesa no es si los mecanismos representaciones de la rana se adecúan a lo que los filósofos esperan de la intencionalidad, sino que si hemos de llamar intencional a la representación de la rana; dicha intencionalidad debe de responder a las demandas genuinas de la rana, no a las del concepto.

¹⁴² En este sentido, si bien la indeterminación teleosemántica proviene de la indeterminación teleofuncional, la demanda de determinación responde a que el rasgo es *cognitivo* no a que el rasgo es funcional *per se*.

¹⁴³ La teoría de Millikan (1984, 2002) incluye un recuento de este hecho ubicuo en el mundo orgánico, a saber, que las estructuras o sistemas biológicos evolucionados suelen tener múltiples funciones propias. Vale la pena presentar su propuesta al respecto. En algunos casos estas funciones son *independientes* entre sí o “disyuntivas” (Millikan, 1984) y como tal, son ejecutadas en distintos momentos. Las funciones “conjuntivas”, en cambio, son ejecutadas simultáneamente. Otro caso de multiplicidad teleofuncional son las funciones “seriales”. Estas funciones están *relacionadas* entre si donde la ejecución propia de una función es una condición normal para la ejecución propia de otra. Finalmente las funciones propias “*focused*” son un tipo de funciones propias seriales (Millikan, 1984, p. 34-38). Éstas hacen referencia al último miembro de dicha serie de funciones, esto es, la última función que dicho mecanismo ejecuta antes de que las funciones seriales diverjan y a partir de ese punto se conviertan en funciones disyuntivas o conjuntivas.

CONCLUSIONES:

El objetivo de esta tesis ha sido el de defender una teoría teleosemántica del contenido de las representaciones mentales (humanas y no humanas) basada en la propuesta de Millikan (1984). La estructura del argumento general está compuesta de tres partes (Walsh, 2002): (1) La explicación de la *función biológica* evolutiva en términos de la *historia causal*. (2) La explicación del *contenido cognitivo* en términos de *función cognitiva*. (3) La explicación de la *función cognitiva* en términos de la *función biológica evolutiva*. La manera en que llegamos de 1 a 3 es mediante los siguientes pasos. En primer lugar, la teoría es exitosa si satisface el criterio naturalista de adecuación:

(CNA): R representa que P es el caso si y sólo si cierta relación C se da entre R y P independientemente de que P sea el caso o no, y donde C no puede a su vez estar caracterizada en términos semánticos o intencionales.

En segundo lugar, la teoría teleosemántica satisface CNA al caracterizar C en términos de la noción etiológica o teleonómica de función biológica:

Definición (df): un caso actual de un rasgo X de un organismo O (miembro de una FRE) en un entorno E tiene la función de producir un efecto del tipo Y , si y sólo si, casos pasados (recientes) de X contribuyeron a la aptitud de los ancestros de O en E al producir Y y fueron seleccionados *por* hacer esta contribución a la aptitud de los ancestros de O frente a variaciones X en E .

El tercer paso (capítulo segundo) consiste en que, dado CNA y (df), generamos la siguiente hipótesis (hipótesis teleosemántica): el pensamiento y la percepción son intencionales en tanto representan cosas y las cosas que representan no tienen que existir para poder ser pensadas o tener percepciones acerca de ellas. Esto es posible en virtud de las relaciones pasadas (ontogenéticas y filogenéticas) entre los estados intencionales y lo que representan:

(HTS): El contenido de una representación caso R producida por un mecanismo cognitivo evolucionado M de un organismo O bajo condiciones C en el entorno E es la de representar C , puesto que representar C causó la selección de M en los ancestros de O en E , en virtud de que representar C mediante R tuvo un efecto positivo en la aptitud de los ancestros de O en E .

Aquí la palabra “representa” parece violentar CNA, pero como vimos en el capítulo segundo, si bien la teleosemántica identifica el contenido de una representación con la función de dicha representación y de los mecanismos que le subyacen, *cómo* es ejecutada esa función varía de teoría en teoría. Millikan (1984, 1993, 2004) reduce la expresión “representa” a la expresión “mapea” y así salva CNA. La idea es que la expresión “mapea” denota un tipo de relación de representación específica:

(RDR): una relación basada en una similitud abstracta, una función matemática de *mapeo* que establece, mediante una regla de *proyección*, un *isomorfismo* abstracto uno-a-uno entre aspectos de la representación o mapa y aspectos de lo representado o mapeado.

Finalmente, si aplicamos de HTS más RDR¹⁴⁴ al contexto de las representaciones mentales que involucran actitudes proposicionales tenemos que¹⁴⁵:

x cree que *P* si y sólo si *x* está en un estado cognitivo *R* que fue *seleccionado* por *mapear* el estado de cosas *P*.

x desea que *P* si y sólo si *x* está en un estado cognitivo *R* que fue *seleccionado* por producir el estado de cosas *P*.

La conclusión es que las propiedades semánticas de las representaciones mentales son propiedades selectivas de los sistemas biocognitivos que las producen y emplean en la

¹⁴⁴ Millikan (1984, 1993, 2004) desarrolla este esquema introduciendo la distinción entre los mecanismos cuya función es la de producir la representación mental (*MP*) y aquellos que cuya función es la de consumirla (*MC*). Para Millikan, *MC* determina tanto el estatus intencional de una representación como su contenido. Puesto que las funciones son efectos que un rasgo se supone que debe producir, la función de una representación mental radica en la manera en que dicha representación contribuye al propósito biológico de *MC*. La función de la representación permite que *MC* alcance su propósito biológico conectando a la conducta con las circunstancias representadas por la representación. Así, las condiciones de verdad o de satisfacción de una representación mental son las circunstancias bajo las cuales la conducta generada realiza el propósito de *MC*. Una consecuencia de esta propuesta es que la determinación del contenido, de las condiciones de verdad o satisfacción no son las condiciones de entrada, *MP*, lo cual permite dar cuenta de la normatividad de la intencionalidad: puesto que el contenido, las condiciones de verdad o satisfacción dependen de la salida conductual, el contenido es la condición bajo la cual el comportamiento resultantes *sería* apropiada, sean o no las circunstancias actuales que causan la representación de ese tipo.

¹⁴⁵ Debemos distinguir entre el teleofuncionalismo, la aplicación del modelo etiológico a las funciones de rasgos biológicos, la teleosemántica, la aplicación del teleofuncionalismo a los rasgos cognitivos representacionales, la teleosemántica modesta, la aplicación del teleofuncionalismo a los rasgos cognitivos representacionales de organismos no humanos, y la teleosemántica inmoles, la aplicación del teleofuncionalismo al contexto cognitivo humano.

producción de conductas específicas. Las propiedades selectivas con las que identificamos las propiedades semánticas corresponden a efectos seleccionados pasados. La originalidad de este naturalismo radica en que la reducción no es física sino biológica y, como tal, está muy lejos de *eliminar* las propiedades (mentales) de nivel superior. Si bien la teleosemántica como propuesta naturalista *sí* implica en un sentido una reducción de propiedades psicológicas a biológicas –en particular de propiedades intencionales a propiedades biofuncionales–, el que la reducción sea a propiedades *biológicas* y no directamente a propiedades físicas permite hablar al mismo tiempo de una *integración* de las propiedades psicológicas al mundo natural. La teleosemántica ofrece así, un modelo ecológico de la cognición donde la historia casual de los sistemas cognitivos determina el contenido de los estados cognitivos. La diferencia entre una reducción física y una biológica es tanto una diferencia *epistémica* como una *metafísica* (el estatus causal de las propiedades selectivas). Es epistémica pues supone que una psicología científica intencional generaría explicaciones en términos ecológico-evolutivos (en lugar de nomológicas). En este sentido, la teleosemántica determina qué tipo de descripciones y modelos resultan relevantes para investigar empíricamente el fenómeno de la intencionalidad.

Una objeción (Fodor, 1990, 1998) contra la idea misma de aplicar el modelo etiológico tanto a funciones biológicas como a funciones cognitivas, humanas y no humanas, revisada en el capítulo tercer, afirma que dicho modelo genera la indeterminación de estas funciones. Esta objeción se disuelve distinguiendo entre la *pluralidad* de funciones propias que corresponden a un sistema, mecanismo, entidad o rasgo biológico y la posible *indeterminación* de dichas funciones. La mayoría de los rasgos funcionales poseen un gran número de efectos cada uno de los cuales hacen algún tipo de contribución a la supervivencia y reproducción donde cada uno es, *prima facie*, un candidato posible a ser la función. Aceptar esta pluralidad no implica aceptar la indeterminación. Sin embargo, dicha objeción ofrece un reto para la versión inmodesta. He defendido que la versión inmodesta puede, *prima facie*, sortear el problema de la indeterminación afirmando que la exigencia de determinación “unívoca” es satisfecha por condiciones relativas al tipo de sistema biológico en cuestión, en este caso, a condiciones propietarias de los sistemas cognitivos de orden superior característicos de la psicología intencional (Millikan, 1989, 1991, 1995, Papineau, 1998, Walsh, 2002). Generalizando esta idea, podemos decir que el grado de determinación es relativo al usuario o consumidor de la

representación.¹⁴⁶ Pero hay otros problemas que tienen que ver con la viabilidad del proyecto teleosemántico y atañen a la necesidad de al desarrollo en una teoría de la representación mental. Uno de ellos es que si bien apelar a relaciones históricas explica que un organismo pensar acerca de cosas con las que no está relacionado en el presente pero con las que mantiene una relación pasada, ¿Qué hay de las cosas que puedo pensar con las que ni siquiera estuve relacionado en el pasado, esto es, que nunca, no ahora ni antes, han existido tales como unicornios o la Fuente de la Juventud? Quizá la respuesta radica en que esos casos son resultado de vincular mentalmente cosas o propiedades de cosas que sí existieron (imaginación). En cualquier caso, es necesario saber más sobre cómo interactúan la filogenia y la ontogenia en la generación de creencias y deseos particulares. Destaca el hecho de que el contenido de muchas de esas actitudes nada tiene que ver con factores ecológicamente relevantes en el sentido estricto de incrementar la aptitud sino lo contrario. Una manera en que el teleosemanticista puede responder a esto es diciendo que los mecanismos relevantes que subyacen a la funcionalidad no tienen que ser necesariamente genéticos ya que algunos rasgos fenotípicos son resultado *además* de la influencia del medio durante la ontogénesis y, procesos ontogenéticos como el aprendizaje, pueden incluir mecanismos de selección. Esto es posible puesto que, desde la perspectiva de la teoría etiológica, lo que hace a un efecto ser una función propia es independiente de si le subyace un proceso genético o no, pues el factor relevante es la intervención de un mecanismo de *selección* que descarta algunos efectos y promueve otros en virtud de algún *beneficio* que tiene para el objeto teleológicamente caracterizado. Esto es material para un trabajo posterior.

¹⁴⁶ Hay otros problemas que no he tocado en esta tesis. Destaca el experimento mental de Pietroski (1992) contra la idea de que el contenido está determinado por la salida conductual más que por la entrada informacional. Dretske (2001) cuestiona la idea de que la mentalidad es constitutivamente histórica y también la idea de que nuestra vida mental está constituida por normas. A sí mismo, Dretske (2006) identifica un problema para la teleosemántica en relación al autoconocimiento. En esta línea, Jackson (2001, 2006) cuestiona la opacidad epistémica que un agente tiene con respecto a sus propios estados mentales y a los de otros, que se sigue de la propuesta teleosemántica.

Apéndice Principales definiciones

Definición de “función propia” de la teoría etiológica histórica de Millikan (1984, p. 28): Donde m sea un miembro de una familia reproductiva establecida R y R tiene el carácter reproductivamente establecido o normal C , m tiene la función F como función propia directa si y sólo si:

1. Ciertos ancestros de m ejecutaron F ;
2. En parte existió una conexión causal directa entre tener el carácter C y ejecutar la función F en el caso de estos ancestros de m , C esta positivamente correlacionada con F sobre cierto conjunto de elemento S que incluyen a estos ancestros y otra cosas que no tengan C ; y
3. Una entre la legítimas explicaciones que puedan darse del hecho de que m existe hace referencia al hecho de que C se correlaciona positivamente con F sobre S , ya sea causando directamente la reproducción de m o explicando porqué R proliferó y por tanto porqué m existe.

Definición de “función propia” de la teoría etiológica de Griffiths (1993, p. 153): Donde i sea un rasgo de sistemas tipo S , una función propia de i en S 's es F si y sólo si una explicación selectiva próxima de la proporción actual diferente a cero de S 's con i debe citar F como un componente en la aptitud conferida por i .

Definición de “función” de la teoría de la historia moderna de Godfrey-Smith (1994, p. 218): La función de m es hacer F si y sólo si:

1. m es un miembro de la familia T ,
2. miembros de la familia T son componentes de sistemas biológicos reales de tipo S ,
3. entre las propiedades copiadas entre miembros de T esta la propiedad o propiedades C , que pueden hacer F ,
4. una razón de que miembros de T como m existan ahora es el hecho de que miembros pasados de T fueron exitosos bajo selección en el pasado reciente, al positivamente contribuir a la aptitud de sistemas del tipo S , y
5. miembros de T fueron seleccionados por qué hicieron F , y llevados a tener C .

Definición de “función” del “análisis funcional” de Cummins (1975, p. 76):

x funciona como un Φ en s (o: la función de x en s es hacer Φ) relativamente a un análisis de A de la capacidad de s 's de hacer Ψ sólo en caso de que x sea capaz de hacer Φ en s y A da cuenta apropiada y adecuadamente de la capacidad de s 's de hacer Ψ mediante, en parte, una apelación a la capacidad de x de hacer Φ en s .

Bibliografías

- Amundson, R., y Lauder, G. V. (1994). "Function without Purpose: the Uses of Causal Role Function in Evolutionary Biology". *Biology and Philosophy*, 9: 443-69. Reimpreso en *The Philosophy of Biology*. Edited by David L. Hull and Michael Ruse. Oxford University Press, 1998. pp. ***
- Beatty, J. (1992). Fitness: Theoretical Contexts. En *Keywords in Evolutionary Biology*. Harvard University Press. Cambridge, Massachusetts London, England. Editado por Lloyd, E. A. & Keller, E. F.
- Bigelow, J., & Pargetter, R. (1987). Functions. En *Function, Selection, and Design* (97-114). State University of New York Press. Editado por David J. Buller.
- Boorse, C. (1977). Wright on functions. *Philosophical Review*, 85, 70-86.
- Brandom, R (1994). *Making it Explicit*. (Harvard University Press).
- (2001). *Articulating Reasons: an Introduction to Inferentialism*. Harvard University Press; New Ed edition (October 15, 2001)
- Buller, D. J. (1999). Natural Teleology. En *Function, Selection, and Design* (1-27). State University of New York Press. Editado por David J. Buller.
- Carnap, R. (1947). *Meaning and Necessity*. University of Chicago Press, Chicago.
- Crane, T. (1995). *The Mechanical Mind. A philosophical introduction to minds, machines and mental representation*. Routledge, London and New York.
- (1998). Intentionality as the mark of the mental. En *Contemporary Issues in the Philosophy of Mind* (229-251). ed. Anthony O'Hear.
- (2001). Intentional objects. En *Ratio*, 14, p. 336-349.
- (2005). Brentano's concept of intentional inexistence. Próximo a salir en un volumen de Austrian Philosophy, editado por Mark Textor (documento de Word). <http://www.homepages.ucl.ac.uk/~uctydtc/Crane%20online%20papers.htm>.
- Cummins, R. (1975). Functional analysis. En *Function, Selection, and Design* (57-83). State University of New York Press. Editado por David J. Buller.
- (1989). *Meaning and mental representation*. Cambridge, MA: MIT Press.
- (2002). Neo-Teleology. En *Functions. New essays in the philosophy of psychology and biology* (157-172). Editado por André Ariew, Robert Cummins, y Mark Perlman. Oxford University Press.
- Danto, A. (1967). *Naturalism*. En: Encyclopedia of Philosophy, ed. Paul Edwards (New York: Macmillan, 1967) vol 5.
- Davidson, D, (1970). 'Mental Events.' En: Lawrence Foster y J. W. Swanson (eds.), *Experience and Theory*. London: Duckworth, reimpreso en Davidson, 2001a, *Essays on Actions and Events*. Oxford: Clarendon Press, 2nd ed.,
- (1974). 'On the Very Idea of a Conceptual Scheme.' En: *Proceedings and Addresses of the American Philosophical Association*, 47, reimpreso en Davidson, 2001b.
- (2001b). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, 2nd ed.
- (2001c). *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press.
- (2004). *Problems of Rationality*. Oxford: Clarendon Press.
- Dawkins, R. (1982). *The extended phenotype*. Oxford: Oxford University Press.
- (1985). *The Blind Watchmaker*. (New York: W. W Norton and Co).
- (1989). *The selfish gene* (2nd ed.). Oxford: Oxford University Press.
- Dennett, D. (1988). Error, evolution and intentionality. En *Sourcebook on the Foundations of Artificial Intelligence*. New Mexico University Press. Y. Wilks and D. Partridge, eds.
- (1995). *Darwin's Dangerous Idea*. A Touchstone Book Published by Simon & Schuster.
- Dewey, J. (1910). *The Influence of Darwin on Philosophy*. New York: Holt, 1910. Reprint ed., Bloomington: Indiana University Press, 1965.
- Dretske, F. (1986). Misrepresentation. En *Belief: Form, Content and Function*, ed. Radu J. Bofdan, Oxford University Press.

- (2001). Norms, History and the Mental. En *Naturalism, Evolution and Mind* (87-104). Editado por Denis M. Walsh. Cambridge University Press.
- Enc, B. (2002). Indeterminacy of Function Attributions. En *Functions. New essays in the philosophy of psychology and biology* (291-313). Editado por André Ariew, Robert Cummins, y Mark Perlman. Oxford University Press.
- Endler, J. A. (1992). Natural Selection: Current Usages. En *Keywords in Evolutionary Biology*. Harvard University Press. Cambridge, Massachusetts London, England. Editado por Lloyd, E. A. & Keller, E. F.
- Fodor, J. (1987). *Psichosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.
- (1990). *A Theory of Content and Other Essays*. A Bradford Book. MIT Press. Cambridge, Massachusetts, London, England.
- (1996). Deconstructing Dennett's Darwin. En *In Critical Condition: Polemical Essays on Cognitive Science and the Philosophy of Mind* (171-187). Cambridge, MA: MIT Press.
- Futuyama, D. (2005). *Evolution*. Sinauer associates, inc. Publishers Sunderland, Massachusetts U.S.A.
- Ghiselin, M. T. (1974). A Radical Solution to the Species Problem. *Systematic Zoology* 23:536-544.
- Godfrey-Smith, P, (1991). "Signal, Decision, Action," *Journal of Philosophy* 88 (1991): 709-722.
- (1992). "Indication and Adaptation," *Synthese* 92 (1992): 283-312.
- (1993). Functions: Consensus without unity. En *Function, Selection, and Design* (285-197). State University of New York Press. Editado por David J. Buller.
- (1994). A modern history theory of functions. En *Function, Selection, and Design* (199-220). State University of New York Press. Editado por David J. Buller.
- (1996). *Complexity and the function of mind in nature*. Cambridge University Press.
- (2001). Three Kinds of Adaptationism. En *Adaptationism and Optimality* (335-356). Cambridge University Press. Editado por Steven Hecht Orzack y Elliott Sober.
- Griffiths, P. E. (1993). Functional analysis and proper function. En *Function, Selection, and Design* (143-158). State University of New York Press. Editado por David J. Buller.
- Griffiths, P y Goode, R, (1995). The Misuse of Sober's Selection for / Selection of Distinction. En *Function, Selection, and Design* (233-242). State University of New York Press. Editado por David J. Buller.
- Hardcastle, V. G. (2002). On the Normativity of Functions. En *Functions. New essays in the philosophy of psychology and biology* (144-156). Editado por André Ariew, Robert Cummins, y Mark Perlman. Oxford University Press.
- Horan, B.L. (1989). Functional explanations in sociobiology. *Biology and Philosophy*, 4, 131-158.
- Hull, D. L. (1978). A Matter of Individuality, *Philosophy of Science* 45:335-360; reprinted in *Conceptual Issues in Evolutionary Biology*, E. Sober (ed.), Cambridge, MA: MIT Press, 1984; and in *The Units of Evolution*, M. Ereshefsky (ed.), Cambridge, MA: MIT Press, 1992.
- (1989). Individuality and selection. En: *The Metaphysics of Evolution* (pp. 89-109). Albany, NY: SUNY Press.
- Hull. D.L. (1998). *The Philosophy of Biology*. Edited by David L. Hull and Michael Ruse. Oxford University Press, 1998.
- Jackson, F. y Braddon-Mitchell, D. (1996). *The Philosophy of Mind and Cognition*. Basil Blackwell.
- Kim, J. (1984). Epiphenomenal and Supervenient Causation, en P. French et. al. ed, *Midwest Studies in Philosophy*, IX, University of Minnesota Press, Minneapolis.
- Kitcher, P. (1993). Function and Design. En *The Philosophy of Biology* (258-279), Edited by David L. Hull and Michael Ruse. Oxford University Press. Editado por David J. Buller.
- Kripke, S. (1984). *Wittgenstein on Rules and Private Language*. Harvard University Press.
- Lettvin, J. Y., Maturana, H. R., McCulloch, W. S. & Pitts, W. H. (1959). *What the frog's eye tells the frog's brain*. En, Proc. Inst. Radio Engr., vol. 47 pp. 1940-1951, reimpresso en "The Mind: Biological Approaches to its Functions" (Editores: William C. Corning, Martín Balaban, 1968, pp. 233-258).
- Loyd, E. (1992). *Keywords in Evolutionary Biology*. Harvard University Press. Cambridge, Massachusetts London, England. Editado por Lloyd, E. A. & Keller, E. F.
- McDowell, J. (1994). *Mind and World* (Cambridge, Mass.: Harvard University Press).

- (1998a). *Mind, Value, and Reality* (Cambridge, Mass.: Harvard University Press, 1998).
- (1998b). *Meaning, Knowledge, and Reality* (Cambridge, Mass.: Harvard University Press, 1998).
- Mills, S. K., y Beatty, J. H. (1979). The propensity interpretation of fitness. *Philosophy of Science*, 46, 263-286.
- Millikan, R. G. (1984). *Language, thought, and other biological categories*. Cambridge, MA: MIT Press.
- (1986). Thoughts without Laws. *Philosophical Review* 95, no. 1: 47-80. Reimpreso en *White Queen psychology and other essays for Alice* (51-82). Cambridge, MA: MIT Press.
- (1991). Speaking Up for Darwin. En *Meaning in Mind. Fodor and his Critics*. Editado por Barry Loewer y Georges Rey. Blackwell, Oxford UK & Cambridge MA.
- (1993). *White Queen psychology and other essays for Alice*. Cambridge, MA: MIT Press.
- (1993a). In defense of proper functions. En *White Queen psychology and other essays for Alice* (13-29). Cambridge, MA: MIT Press.
- (1993b). Biosemantics. En *White Queen psychology and other essays for Alice* (83-101). Cambridge, MA: MIT Press.
- (1993c). Thoughts without Laws. En *White Queen psychology and other essays for Alice* (51-82). Cambridge, MA: MIT Press.
- (1993d). Explanation in biopsychology. En *White Queen psychology and other essays for Alice* (171-192). Cambridge, MA: MIT Press.
- (1996). *On Swampkinds*. *Mind and Language* 11.1 (March 1996), pp. 103-117.
- (1999). An ambiguity in the notion "function". En *Function, selection and design* (115-121) Editado por David J. Buller. Albany, NY: SUNY Press.
- (2002). Biofunctions: two paradigms. En *Functions: new essays in the philosophy of psychology and biology* (113- 143). Editado por André Ariew, Robert Cummins, and Mark Perlman. Oxford: Oxford University.
- (2004). Varieties of Meaning. The 2002 Jean Nicod Lectures. MIT Press.
- Mitchell. S. D. (1993). Dispositions or etiologies? A comment on Bigelow and Pargetter. *Journal of Philosophy*, 90, 249-259.
- Neander, K. (1991). The teleological notion of "function". En *Function, Selection, and Design* (123-141). State University of New York Press. Editado por David J. Buller.
- (1995). *Misrepresenting and malfunctioning*. *Philosophical Studies*, 79, 109-141.
- Papineau, D. (1987). *Reality and representation*. (New York: Blackwell).
- (1993). *Philosophical Naturalism*. (Oxford: Blackwell).
- (1998). Teleosemantics and Indeterminacy. En *Australasian Journal of Philosophy*, 76: 1-14.
- Perlman, M. (2002). Pagan Teleology: Adaptational Role and the Philosophy of Mind. En *Functions. New essays in the philosophy of psychology and biology* (265-290). Editado por André Ariew, Robert Cummins, y Mark Perlman. Oxford University Press.
- Popper, K.(1959). *The Logic of Scientific Discovery*. Routledge; 1 edition (March 29, 2002).
- Price, C. (1998). *Determinate Functions*. *Noús*, 32, pp. 54-75.
- Segal, G. y Sober, E. (1991). The Causal Efficacy of Content. *Philosophical Studies*. 63: 1-30.
- Sellars, W. (1956). *Empiricism and the Philosophy of Mind* Republished as a book with an introduction by Richard Rorty, and a study guide by Robert Brandom (Cambridge, Mass.: Harvard University Press, 1997).
- Shapiro, L. A. (1992). Darwin and disjunction: Optimal foraging theory and univocal assignments of content. *Proceedings of the Philosophy of Science Association*, 1, 469-480.
- Sober, E. (1984). *The nature of selection*. Cambridge, MA: MIT Press.
- (2001). The Principle of Conservatism in Cognitive Ethology. (2001). En *Naturalism, Evolution and Mind* (87-104). Editado por Denis M. Walsh. Cambridge University Press.
- Stalnaker, R. C. (1999). *Context and Content*. Oxford University Press.
- Sterelny, K, (2003). *Thought in a Hostile World*. (Blackwell).
- y Griffiths, P. (1998). *Sex and Death: An Introduction to Philosophy of Biology*. (University of Chicago Press).

- Stich, S. (1994). What Is a Theory of Mental Representation?. En *Mental Representation A Reader* (347-364). Editado por Stephen P. Stich y Ted A. Warfield. Blackwell, Oxford UK y Cambridge USA.
- Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, 20, 410-433.
- Tye, M. (1992). Naturalism and the Mental. En *Mind*, New Series, Vol. 101, No. 403 (Jul., 1992), pp. 421-411).
- Wade, M. J. (1992). Heritability: Historical Perspectives. En *Keywords in Evolutionary Biology*. Harvard University Press. Cambridge, Massachusetts London, England. Editado por Lloyd, E. A. & Keller, E. F.
- Walsh, D. M., & Ariew, A. (1996). A taxonomy of functions. En *Function, Selection, and Design* (257-279). State University of New York Press. Albany, NY: SUNY Press. Editado por David J. Buller.
- Walsh, D. M. (2002). Brentano's Chestnuts. En *Functions. New essays in the philosophy of psychology and biology* (314-337). Editado por André Ariew, Robert Cummins, y Mark Perlman. Oxford University Press.
- West-Eberhard, M. J. (1992). Adaptation: Current Usages. En *Keywords in Evolutionary Biology*. Harvard University Press. Cambridge, Massachusetts London, England. Editado por Lloyd, E. A. & Keller, E. F.
- Wimsatt, W. (1972). *Teleology and the logical structure of function statements*. *Studies in History and Philosophy of Science*, 3, 1-80.
- (2002). Functional Organization, Analogy, and Inference. En *Functions. New essays in the philosophy of psychology and biology* (173-221). Editado por André Ariew, Robert Cummins, y Mark Perlman. Oxford University Press.
- Wittgenstein, L. (1953) *Investigaciones filosóficas*. Editorial Crítica-UNAM (****).
- Wright, L. (1973). Functions. En *Function, Selection, and Design* (29-55). State University of New York Press. Albany, NY: SUNY Press.
- (1976). *Teleological explanation*. Berkeley: University of California Press.